Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research

Sally S. Dickerson and Margaret E. Kemeny
University of California, Los Angeles

This meta-analysis reviews 208 laboratory studies of acute psychological stressors and tests a theoretical model delineating conditions capable of eliciting cortisol responses. Psychological stressors increased cortisol levels; however, effects varied widely across tasks. Consistent with the theoretical model, motivated performance tasks elicited cortisol responses if they were uncontrollable or characterized by social-evaluative threat (task performance could be negatively judged by others), when methodological factors and other stressor characteristics were controlled for. Tasks containing both uncontrollable and social-evaluative elements were associated with the largest cortisol and adrenocorticotropin hormone changes and the longest times to recovery. These findings are consistent with the animal literature on the physiological effects of uncontrollable social threat and contradict the belief that cortisol is responsive to all types of stressors.

An extensive animal and human literature documents that psychological factors can influence the hypothalamic–pituitary–adrenocortical (HPA) axis, which regulates the release of cortisol, an important hormone associated with psychological, physiological, and physical health functioning. Over the past half century, hundreds of studies have specifically focused on the effects of psychological stressors on cortisol activation. Despite the magnitude of this research enterprise, only two broad conclusions can be drawn from this literature as a whole. First, like physical stressors (e.g., electric shock, prolonged exercise), psychological stressors are indeed capable of activating the HPA axis; a number of studies have reported that laboratory tasks such as public speaking or mental arithmetic can increase cortisol levels (e.g., Kirschbaum, Pirke, & Hellhammer, 1993). Second, the effects of psychological stressors on this physiological system are highly variable. Many studies have failed to find cortisol changes (e.g., Manuck, Cohen, Rabin, & Muldoon, 1991), and recent narrative reviews have highlighted the inconsistent effects of psychological stressors on cortisol activity (e.g., Biondi & Picardi, 1999). The tremendous heterogeneity in the literature suggests that all types of negative situations may not uniformly trigger cortisol changes (Mason, 1968). Essential elements, present only in contexts that elicit cortisol responses, have yet to be clearly delineated.

What is it, then, about certain conditions that would make them capable of inducing a cortisol response? For decades, this fundamental question has generated a spectrum of hypotheses. Hans Selye (1956) argued that the stress response, which includes HPA activation, was nonspecific: All stressors, whether physical or psychological, would elicit the same physiological reaction. Others have concluded from the early work investigating the effects of severe traumatic experiences on cortisol activity (e.g., electric shock, injury) that only extreme or prolonged stressful conditions trigger cortisol elevations. Some have focused on the specific characteristics of the stressor, hypothesizing that contexts that are novel (Rose, 1980), unpredictable (Mason, 1968), uncontrollable (Henry & Grim, 1990; Sapolsky, 1993), or threatening, with the potential for harm or loss (Blascovich & Tomaka, 1996; Dienstbier, 1989), would be most likely to activate this system. Although a number of hypotheses have been offered, many have never been empirically tested, and in other cases, the evidence is not as conclusive as popular wisdom may suggest. For example, although uncontrollable contexts are commonly thought to elicit cortisol responses, the support for this association stems primarily from nonhuman animal studies; primates or other animals that had control over electric shocks or blasts of noise showed attenuated cortisol responses compared with “yoked” animals that received identical stimuli without control (Davis et al., 1977; Dess, Linwick, Patterson, Overmier, & Levine, 1983; Hanson, Larson, & Snowdon, 1976; Swenson & Vogel, 1983; Weiss, 1971). However, there is surprisingly little empirical evidence for a relationship...
between acute, uncontrollable conditions and cortisol activation in humans (for review, see Peters et al., 1998).

Although it remains unclear whether psychological stressors with specific characteristics preferentially elicit cortisol changes in humans, research in animals supports the premise that there could be stressor-specific pathways to cortisol activation. For example, exposing animals to distinct types of physical, or systemic, stressors (e.g., heat, shock) can lead to different effects on the HPA system (Weiner, 1992). Systemic stressors have been differentiated from psychological (neurogenic) stressors in terms of their neural correlates and downstream physiological effects, including activation of components of the HPA system (e.g., Sawchenko & Ericsson, 2000). In addition, distinctive physiological correlates have been found for different stress-relevant behavioral patterns in animals (e.g., fighting, fleeing, submitting; Weiner, 1992). However, little is known about the differential impact of types of psychological stressors on the cortisol system in humans.

Elucidating the conditions in which psychological stressors activate the cortisol system has several important implications, not only for the field of psychobiology, but for broader psychological theory and research as well. First, psychological stressors affect physiology by activating specific cognitive and affective processes and their central nervous system underpinnings. The thalamus and frontal lobes (e.g., prefrontal cortex) first integrate sensory information and evaluate or appraise the significance or meaning of environmental stimuli. These cognitive appraisals can lead to the generation of emotional responses via extensive connections from the prefrontal cortex to the limbic system (e.g., the amygdala and hippocampus). The limbic structures, which connect to the hypothalamus, serve as a primary pathway for activating the HPA axis (see Feldman, Conforti, & Weidenfeld, 1995, or Lovallo, 1997, for reviews on central nervous system inputs to the HPA system).

Activation of the HPA axis is initiated by the hypothalamic release of corticotropin releasing hormone (CRH), which stimulates the anterior pituitary to secrete adrenocorticotropin hormone (ACTH), which in turn triggers the adrenal cortex to release cortisol into the bloodstream (for review, see Lovallo & Thomas, 2000; Sapolsky, Romero, & Munch, 2000). In particular, the paraventricular nucleus of the hypothalamus, a cell group that expresses CRH, is believed to play a key role in translating the differentiated neural activation patterns generated under specific environmental conditions into specific patterns of physiological and behavioral responses, including activation of the HPA axis (Sawchenko & Ericsson, 2000). Research documenting an association between specific stressors and cortisol responses might inform an understanding of the links between the cognitive and affective responses associated with specific stressful circumstances, the neural substrates of these responses, and activation of the HPA system.

Second, the HPA axis is vital for supporting normal physiological functions and regulating other systems. Cortisol plays a critical role in metabolism by mobilizing energy resources to provide “fuel” for the body. This is primarily accomplished by elevating blood glucose levels (by stimulating the conversion of amino acids and other substrates to glucose in the liver, and promoting the breakdown of protein and fat stores in the tissue); the net result is the release of energy reserves that allow adequate metabolic functioning. Cortisol is an important regulator of other physiological systems. For example, cortisol can inhibit many aspects of immune system functioning. It can be considered the body’s own natural anti-inflammatory because it can preferentially inhibit proteins that play a central role in regulating inflammation. Cortisol also has permissive effects, which allow other physiological systems to function effectively. For example, certain levels of this hormone are necessary for the catecholamines and other sympathetic products to exert effects on the cardiovascular system (e.g., induce vasoconstriction, increase heart rate). Therefore, the specific conditions that elevate cortisol levels also have the potential to influence the variety of critical physiological processes that can be affected by HPA activity.

Third, when the HPA system is activated, it is associated with important cognitive and affective processes and is thought to have implications for health and disease. Heightened HPA activity has been associated with depressive symptomology (E. S. Brown & Suppes, 1998; Heim & Nemeroff, 1999) and can have effects on memory (e.g., Buchanan & Lello, 2001; Kirschbaum, Wolf, May, Wippich, & Hellhammer, 1996; Lupien et al., 1997). Prolonged cortisol activation (produced by frequent exposure to stressors or by failing to shut down this response after stressor termination) is associated with a number of negative biological and health effects, including suppression of aspects of the immune system (e.g., decreased lymphocyte proliferation and cytokine production); damage to hippocampal neurons; and the development and/or progression of certain chronic diseases, such as diabetes and hypertension (Boomershine, Wang, & Zwilling, 2001; McEwen, 1998). Therefore, determining whether specific stressors activate the cortisol system could delineate the conditions capable of contributing to the onset or exacerbation of certain health outcomes.

A meta-analytic review of the studies examining acute stressors and cortisol responses provides an opportunity to address the theoretical debate on the elicitors of this system as well as helps to explain the tremendous variability in this literature. Although several narrative reviews have provided overviews of topics relevant to acute stressors and cortisol responses (e.g., Biondi & Picardi, 1999; Kirschbaum & Hellhammer, 1994; Lovallo & Thomas, 2000; Stansbury & Gunnar, 1994), in most cases, their primary purpose was not to address the specific conditions that elicit cortisol activation; there has not been a quantitative review that evaluates this literature as a whole. This type of research synthesis could systematically assess the characteristics that predict cortisol responses across a comprehensive set of acute psychological stressor studies.

The purpose of this meta-analysis is twofold: (a) to present a quantitative review of 208 acute psychological laboratory stressor studies that assess cortisol as an outcome, and (b) to test a theoretical framework that delineates the conditions most likely to elicit cortisol responses. Drawing on theory and empirical research in animals and humans, we propose that uncontrollable threats to the goal of maintaining the “social self” would trigger reliable and substantial cortisol changes. In the subsequent sections, we outline the theoretical rationale and empirical support for specifically linking uncontrollability and threats to the social self (e.g., social-evaluative conditions) to cortisol activation. Then, through meta-analysis, we test whether these specific threats influence both the overall magnitude of the cortisol response and patterns of recovery (i.e., the degree to which elevations persist after the stressor ends).
In addition, using a subset of the studies, we determine the effects of these specific stressors on ACTH secretion.

Recovery may be a critical issue because failing to shut down the cortisol system after provocation could lead to greater overall exposure to this hormone, which could have deleterious effects on health (e.g., Linden, Earle, Gerin, & Christenfeld, 1997; McEwen, 1998; Sapolsky et al., 2000). Despite the potential implications, very few studies have examined factors that contribute to delayed recovery (cf. Earle, Linden, & Weinberg, 1999; Matthews, Gump, & Owens, 2001; Roy, Kirschbaum, & Steptoe, 2001); the research synthesis provides an opportunity to examine whether uncontrollable conditions that threaten the social self influence recovery processes.

In addition, whereas some studies have reported strong correlations between ACTH and cortisol responses to stressors (e.g., Kirschbaum, Kudielka, Gaab, Schommer, & Hellhammer, 1999), others have found dissociations between these two indicators of HPA activity (e.g., Cacioppo et al., 1995; van der Pompe, Antoni, & Heijnen, 1996). By examining the subsample of studies that also assessed ACTH secretion, we can determine, through meta-analysis, whether patterns of stress-induced ACTH and cortisol changes occur in concert, and whether social-evaluative, uncontrollable conditions are associated with heightened activity across multiple levels of the HPA system.

**Theoretical Model**

**Activation of the HPA System**

Although the HPA axis is often considered a general system that can be activated in response to a variety of negative situations (or in response to all types of ”stress”; e.g., Selye, 1956), an alternative perspective argues that cortisol is released in response to a more circumscribed set of eliciting conditions. According to Weiner (1992),

The behavioral and physiological responses of the organism to a particular stressful experience, unless overwhelming, are very specific (not general). . . . [Hormone] secretion subserves the metabolic and behavioral requirements of the organism in its efforts to survive and overcome danger and challenge. (p. 243)

It has been argued that the cortisol system is activated under conditions in which central goals are threatened or impediments to attaining desired goals are encountered (Blascovich & Tomaka, 1996; Carver & Scheier, 1999; Dienstbier, 1989; Lazarus & Folkman, 1984). This motivational perspective assumes that “without goal commitment, there would be nothing of adaptational importance at stake in an encounter to arouse a stress reaction” (Lazarus, 1999, p. 76). Although a variety of circumstances can elicit negative feeling states, only those that threaten a central goal are considered capable of triggering this particular physiological response.

For example, threats to the goal of physical self-preservation (i.e., survival, safety) can elicit cortisol changes. Responses to these survival threats include activation of the HPA system, presumably because cortisol mobilizes energy resources and modulates other physiological systems to effectively respond to the short-term metabolic demands of the threat (Lovallo & Thomas, 2000; Sapolsky et al., 2000). The mobilization of energy-relevant systems and the reduction in restorative systems are adaptive when behavioral output is needed to reduce the threat. Although threats to the central goal of physical self-preservation are regarded as the prototypical conditions that trigger HPA activation (Sapolsky et al., 2000), a growing animal and human literature indicates that threats to other central goals could activate this system as well.

**Social Self-Preservation Theory**

Similar to the motive of preserving the physical self, we propose that the motive to maintain and preserve the social self is supported by specific biological processes that include HPA activation (Dickerson, Gruenewald, & Kemeny, in press; Kemeny, Gruenewald, & Dickerson, 2004). This system, which we term the social self-preservation system, monitors the environment for threats to one’s social esteem or social status and coordinates psychological, physiological, and behavioral responses to cope with such threats. Responses to these threats include increases in negative self-evaluations (i.e., negative self-related cognitions and emotions), increases in cortisol, and changes in other physiological parameters. The magnitude of these responses depends on the intensity of the threat, its context, and the presence of vulnerability and protective factors in the individual and social environment.

The social self reflects one’s social value, esteem, and status and is largely based on others’ perceptions of one’s worth (de Waal, 1989; Gilbert, 1997). Individuals who possess qualities that are valued by the group are positively regarded, respected, and esteemed by others and have high social standing. Conversely, those that lack these valued attributes or have undesired characteristics receive signals of rejection or disinterest from group members and are lower on the social hierarchy. The quality and valence of the social self is formed through these social assessments.

Humans are driven to preserve the social self and are vigilant to threats that may jeopardize their social esteem or status. A number of theories propose the existence of motives and goals similar to social self-preservation, including the need for positive self-presentation, social status, and positive self-regard (Allport, 1937; Baumeister & Leary, 1995; Bowlby, 1969; James, 1890/1950; Leary & Kowalski, 1990; Maslow, 1987; McClelland, 1984; Taylor & Brown, 1988). This motivational domain can also be seen across a wide variety of species; nonhuman primates and other animals have developed adaptive psychobiological responses to threats to social status in hierarchies (Sapolsky, 1993). Although the social self-preservation system takes a more complex form in humans, we believe that the phylogenetic roots for this system can be observed in other social animals.

Primates have complex social systems in which hierarchies emerge, providing a salient marker of rank or social status relative to the others in the troop. The hierarchy is maintained through aggressive displays by dominants that are reciprocated by submissive behavior by subordinates, which continually reinforces the latter’s lower status. Subordinate, low-rank animals consistently have higher levels of HPA activation when compared with those of higher rank (e.g., Sapolsky, 1993), and the frequency of submissive display behavior correlates with cortisol activity (Shively, Laber-Laird, & Anton, 1997). Uncontrollable contexts can augment this effect; cortisol levels of subordinate primates are particularly elevated when conditions are uncontrollable or unstable (Sapolsky, 1993). Studies that have manipulated social position
in the hierarchy demonstrate the acute effects of lowered social status. Shifting from dominant to subordinate rank is associated with concomitant increases in cortisol levels (ShivELY et al., 1997).

After an antagonistic encounter, an animal that is defeated and drops in social rank shows greater cortisol activity compared with the victor (e.g., KOLLACK-Walker, Watson, & Akil, 1997; PICH et al., 1993).

Taken together, these studies indicate that acute or chronic threats to social status can lead to increases in cortisol activity in primates and other animals, particularly when conditions are uncontrollable; comparable forms of social threat could trigger cortisol changes in humans as well. Many theorists have argued that social hierarchies exist within human groups across cultures; like nonhuman primates, humans organize their social groups so that some individuals are more highly regarded and have higher status relative to others (e.g., Fiske, 1992). Status can be conferred in humans, as in lower animals, through power, dominance, and ability to influence by means of a threat-based, agonistic system that relates to access to resources. However, status is more commonly obtained in humans through hedonic processes that relate to respect, social esteem, acceptance, and positive social attention (de Waal, 1989; Gilbert, 1997; Gilbert & Trower, 1990). Although threats to social esteem and dominance may often overlap (e.g., situations that threaten one’s social esteem, respect, and/or acceptance could also threaten how much influence or dominance one has), they do not always co-occur (for discussion, see Leary & Baumeister, 2000; Leary, Cottrell, & Phillips, 2001). We argue that threats to social esteem, respect, and acceptance, either in or outside the context of dominance, can activate the HPA system, causing the release of cortisol. These threats may provide the human analogue to social status threats in animals and cause activation of this system (Gilbert, 1997).

Threats to the goal of preserving the social self in humans would include situations that require displays of valued attributes or skills in the presence of others, as demonstrating a lack of these qualities (through poor performance or failure) could lead to a loss of social esteem and/or social status. These social-evaluative conditions are characterized by potential or explicit social rejection and therefore could have implications for other social goals. For example, the need to preserve the social self can overlap with the need for interpersonal belonging, as failing to maintain social esteem may decrease the likelihood of forming close personal ties with others (i.e., friendship).

Conditions that threaten the social self can elicit negative self-evaluations, as assessments of how we are viewed by others fundamentally affects how we see ourselves (e.g., Baumeister, 1998; Cooley, 1902/1983; Hardin & Higgins, 1996; Mead, 1934). Several lines of research indicate that social evaluation can lead to changes in self-esteem and self-related emotion (e.g., shame, embarrassment). Leary and colleagues (Leary & Baumeister, 2000; Leary, Tambor, Terdal, & Downs, 1995) have proposed that self-esteem and self-related emotion are driven by the degree to which others are accepting or rejecting of the self; perceived threats to social esteem or acceptance lead to increases in negative self-evaluative states (Gilbert, 1997; Leary et al., 1995, 2001). The presence and/or evaluation of others can lead to social comparisons (e.g., Swallow & Kuiper, 1988; Taylor, Neter, & Wayment, 1995) or self-awareness (Carver & Scheier, 1981; Pyszczynski & Greenberg, 1987), which can also initiate negative self-evaluative processes. The resulting negative self-related states may mediate the effects of threats to the social self on physiological systems (Dickerson, Gruenewald, & Kemeny, in press).

**Social-Evaluative Threat**

We hypothesize that contexts characterized by social evaluation would elicit a significant cortisol response, as a result of the salient threat that it would pose to the goal of maintaining the social self. Social-evaluative threat occurs when an important aspect of the self-identity is or could be negatively judged by others. We propose that social-evaluative threat is most likely to occur when failure or poor performance could reveal lack of a valued trait or ability. Whereas certain characteristics may be important under circumscribed conditions or among certain groups (e.g., athletic ability), others, such as intelligence or competence, are considered core attributes that are widely valued across diverse domains (e.g., Crocker & Wolfe, 2001; Kirkpatrick & Ellis, 2001; Leary & Baumeister, 2000). In the laboratory, motivated performance situations provide conditions in which these core attributes are vulnerable, because they are active performance tasks that require or demand overt or cognitive responses and have the potential for evaluation (e.g., mental arithmetic, speech task; Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996). Because social-evaluative threat creates the potential for loss of social esteem, it could heighten the stakes for failure in a performance-relevant situation (Seta & Seta, 1995).

There is emerging empirical evidence for this relationship between social-evaluative threat and cortisol responses. One experimental study has demonstrated that social exclusion increases cortisol levels (Stroud et al., 2000). Individuals with characteristics that would make them particularly sensitive to social evaluation show exaggerated cortisol responses to acute stressors. For example, children low in social competence show greater cortisol changes to a social-evaluative stressor (peer self-presentation task; Schmidt et al., 1999), and several studies have demonstrated that individuals with low-self esteem show elevated cortisol responses to laboratory stress tasks (Kirschbaum, Pruessner, et al., 1995; Pruessner, Hellhammer, & Kirschbaum, 1999; Seeman, Berkman, et al., 1995). Although these studies demonstrate that those who might be vulnerable to social evaluation show heightened cortisol activity, a more direct test of our hypothesis would compare contexts that vary in social-evaluative threat to elucidate the specific conditions that trigger the cortisol response.

**Uncontrollability**

The animal literature suggests that cortisol responses to social status threats can be heightened when conditions are uncontrollable. Uncontrollability could impede the process of attaining social esteem and/or social status in humans, which could increase activation of the cortisol system when the social self is threatened. In uncontrollable conditions, a behavioral response cannot affect an outcome (Averill, 1973; Levine & Ursin, 1991; Thompson, 1981; Weiner, 1992). This creates a context of forced failure, in which participants are unable to avoid negative consequences or cannot succeed despite their best efforts. Because there is nothing that can be done to change the situation, uncontrollability could greatly amplify the goal threat, which therefore could lead to exaggerated
cortisol responses. Many have posited that uncontrollability is the specific stressor dimension that triggers cortisol activation (e.g., Henry & Grim, 1990; Sapolsky, 1993), and a number of nonhuman animal studies support this contention.

There has been surprisingly little empirical work in humans that has examined the relationship between acute, uncontrollable psychological situations and cortisol responses. Several studies offer indirect support for an association between cortisol and uncontrollability. Lovato, Pincomb, Brackett, and Wilson (1990) found that when participants failed to complete a goal-relevant task (reaction time task with punishment), whereas no changes were found after a more controllable appetitive task (reaction time task with monetary incentive). Lundberg and Frankenhaeuser (1980) found a positive relationship between cortisol levels and subjective responses to five experimental tasks completed on separate days, and found that cortisol excretion loaded on a distress factor. Because uncontrollable tasks are thought to induce effort with distress (Frankenhaeuser, 1991), this factor analysis has been interpreted as support for the association between uncontrollable conditions and cortisol responses. When uncontrollability has been experimentally manipulated, results have been mixed; some studies have found support for the cortisol–uncontrollability link (Breier, 1989; Croes, Merz, & Netter, 1993; Peters et al., 1998), whereas others have not (Bohlin, Eliasson, Hjemdahl, Klein, & Frankenhaeuser, 1986; Bohlin, Eliasson, Hjemdahl, Klein, Fredrikson, & Frankenhaeuser, 1986; Steptoe, Fieldman, Evans, & Perry, 1993). Despite strong support from the animal literature, the evidence for whether cortisol is released specifically in response to uncontrollable situations in humans is inconclusive; therefore, this remains a key theoretical question.

Uncontrollable conditions may have greater effects on cortisol responses when the outcome of the situation affects an important domain or impedes progress toward a salient goal. In other words, the combination of threat to a central goal and an inability to overcome that threat may lead to substantial HPA activation. For example, in many of the animal studies, procedures (e.g., electric shock) directly threatened the physical integrity of the animal, suggesting that the uncontrollability occurred in a goal-relevant context. In humans, motivated performance situations with social-evaluative threat could provide one set of conditions in which an important goal is threatened.

The Research Synthesis

A research synthesis provides the opportunity to look across all of the acute laboratory stressor studies to examine the specific experimental conditions that elicit cortisol responses. This approach allows us to test one component of our theoretical model: that uncontrollable threats to the goal of maintaining the social self trigger cortisol activation. Specifically, we hypothesized that uncontrollable motivated performance situations with social-evaluative threat (goal-relevant conditions in which others could observe the performance and failure is the likely outcome) would lead to larger cortisol elevations than stressors without these characteristics.

A meta-analytic framework could be particularly germane for clarifying the relationship between acute psychological stressors and cortisol responses because methodological and procedural factors could contribute to the inconsistent results in the literature. For example, cortisol levels show a circadian rhythm, in which levels increase dramatically on awakening and gradually decrease throughout the day, reaching the lowest levels late in the evening; therefore, the time of day that the study was conducted could be a confounding variable. The timing of cortisol assessment from stressor onset could lead some studies to miss changes in cortisol, as there is a time lapse between the onset of an acute stressor and the peak cortisol response. The time of day, timing of assessment, and other methodological variables could preclude eliciting and/or capturing cortisol responses to acute laboratory stressors (e.g., Kirschbaum & Hellhammer, 1994; Lovato & Thomas, 2000; Mason, 1968). These factors could not only obscure possible relationships between psychological stressors and cortisol activation, but also greatly hinder the ability to compare results across studies. Meta-analysis provides the opportunity to first identify the methodological factors associated with cortisol responses and then statistically control for them, potentially elucidating previously masked relationships between specific eliciting conditions and cortisol activation.

We limited the research synthesis to studies that assessed cortisol responses in healthy adults because the pathophysiology that accompanies many psychological or physical disorders could reduce the interpretability of meta-analytic results. Furthermore, we examined acute stressors in a laboratory setting because it permits greater standardization of the stressor task and control over confounding variables, and allows for direct causal inferences between stressful conditions and cortisol changes. Therefore, this meta-analysis does not address stressful circumstances of long duration and those that cannot be modeled in a laboratory. Although it is clear that individual difference factors and appraisal processes are important for understanding individual variation in response to stressors, a solid conceptual understanding of the contexts that reliably trigger cortisol responses across individuals paves the way for the second generation of research questions, leading to more focused research on those factors that may be particularly relevant for understanding individual variability in cortisol responsivity.

A sophisticated quantitative assessment of the conditions associated with cortisol changes can address several fundamental questions regarding the elicitors of this system. First, studies conducted in this area have used a number of different stressor tasks that can vary immensely along dimensions such as duration, controllability, and relevance for important goals. We capitalize on this heterogeneity and include in the meta-analysis the broad range of tasks commonly used in the stress reactivity literature. Thus, the meta-analysis provides an estimate of the magnitude of the cortisol response to acute psychological stressors across all of the tasks used in this area, and determines whether this activation is nonspecific with regard to the nature of the stressor. Second, by coding the stressor category and characteristics of the tasks, we systematically evaluate the conditions associated with cortisol responses, and specifically test our theoretical model that uncontrollable threats to maintaining the social self are capable of triggering relatively large changes in this system. Third, because the process of recovery may have different predictors than the overall magnitude of the cortisol response (e.g., Linden et al., 1997), we also determine whether these conditions affect the degree to which cortisol elevations may persist after stressor termination. Finally, we can examine whether social-evaluative, uncontrollable condi-
tions also lead to heightened ACTH responses, demonstrating consistent effects across levels of the HPA system.

Method

Selection of Studies

The research synthesis was limited to studies that used acute psychological laboratory stressor tasks. We defined acute laboratory stressors as tasks that lasted 1 hr or less and did not serve a function outside the laboratory setting; this excluded extended stressor challenges, chronic stressor studies (e.g., caregiving), and naturalistic stressors (e.g., class examination). Psychological stressors were defined as non-metabolically demanding tasks, which excluded physical stressors (e.g., exercise), physical–psychological stressor combinations (e.g., cold pressor, hand-grip), and studies that involved a biological challenge or placebo injection (e.g., CRH, caffeine). This definition of acute psychological laboratory stressors included tasks typically used in stress reactivity research, such as cognitive tasks, public speaking tasks, marital conflict interactions, noise exposure, and emotion induction procedures (e.g., provocative films, music).

In addition to the type of stressor used, three other criteria determined inclusion in the meta-analysis. First, the study must have used healthy adult participants. This excluded studies in which recruitment was based on a physical or psychological diagnosis or a stressful experience (e.g., diabetes, depression, bereavement), studies that examined stress reactivity in children and/or adolescents (mean age was under 18 or the age range included participants under 18), and studies in which participants did not have full cognitive capacities (e.g., under hypnosis). Second, salivary or plasma cortisol must have been assessed before the onset of the stressor and again during or after the stressor. Studies that used urinary cortisol measures were excluded because this type of collection less adequately captures short-term changes in response to acute stressors (Lovatto & Thomas, 2000). Finally, the study must have been published in an English-language scientific journal or abstracted in Dissertation Abstracts International.

To identify relevant studies, computerized searches (on MEDLINE and PsycINFO) were performed, combining the truncated key words cortisol, HPA, neuroendocrine, hydrocortisone, psychoneuroimmunology, psychoneuroendocrinology, and psychoneuroendocrinology with stress reactivity, acute stress, laboratory stress, experimental stress, psychological stress, mood induction, and emotion. The search included all studies published through 2002 (and indexed by February 2003). This search generated 6,891 articles (4,395 on MEDLINE and 2,496 on PsycINFO; although there was overlap between databases). These abstracts were reviewed, and all articles that could not be conclusively excluded on the basis of the abstract were retrieved to determine eligibility. In addition, the reference lists of all qualifying articles and several recent reviews (Biondi & Picardi, 1999; Kirschbaum & Hellhammer, 1994) were examined for additional studies not identified through the computer search.

The final sample included 208 studies reported in 187 journal articles. These qualifying studies are marked with a single asterisk in the reference section.

Coded Variables

We coded all of the studies that met the inclusion criteria to assess characteristics of the participants, method, and stressor. A trained judge coded a randomly selected 15% of the studies to assess interrater reliability (see reliabilities below). Reliabilities were calculated with the intraclass correlation ($r_I$) for continuous variables and kappa for categorical variables (Orwin, 1994).

Participant Characteristics

For each study, we coded (a) the number of participants ($n = .97$), (b) the mean age of the participants ($r_I = .99$), (c) the gender composition of participants (coded as percent male; $r_I = 1.00$), and (d) any health behavior or disease criteria on which participants were screened and excluded (e.g., depression, autoimmune disease; $r = .93$).

Methodological Characteristics

Time of day ($r_I = .99$). Because cortisol has a circadian rhythm and levels vary dramatically over the course of the day, the time of day at which the study was conducted could be an important moderating variable. Studies in which the stressor task was initiated before 12 p.m. were coded as morning studies, and those with stressor tasks that began after 12 p.m. were coded as afternoon studies. When possible, the time of stressor onset, rather than time of arrival at the laboratory, was used to determine the time of day code. When the time of day was not constant across participants, too broad to fit into a circumscribed code (e.g., between 9 a.m. and 4 p.m.), or not reported, the time of day was coded as unassessed.

Timing of cortisol assessment ($r_I = .98$). There is a delay in detecting elevations in cortisol from the onset of stressful experience, as it takes time to activate the HPA axis. However, it is unclear exactly how long this lag is in the context of acute stressors, so timing of assessment could be an important factor in capturing cortisol responses. Timing of assessment was coded as minutes from stressor onset and treated as a continuous variable.

Method of cortisol assessment ($r_I = .91$). In the context of acute stressors, cortisol can be reliably assessed in either plasma or saliva. Plasma samples reflect levels of cortisol bound to protein as well as biologically active free cortisol (unbound). However, salivary samples reflect only the levels of free cortisol. Because salivary and plasma samples assess different cortisol fractions and reactivity to venipuncture could affect cortisol values, method of assessment was coded for each study.

Other methodological features. Other methodological features of the study could affect cortisol responses by reducing error variance. We coded two aspects of methodological rigor:

1. In investigations where an excluded type of stressor (e.g., metabolically demanding task) and an acute psychological stressor were performed on the same day, studies were excluded when (a) the acute psychological laboratory stressor followed the physical stressor, or (b) the acute psychological laboratory stressor occurred before the excluded type of stress task, but cortisol was assessed after the onset of the excluded stressor. In investigations where an excluded type of stress task and an acute psychological stressor were conducted on separate days, only the psychological stressor data were included.

2. Emotion induction studies were included in the meta-analysis when they provided evidence that a negative mood was successfully induced. This included using a validated emotion-induction stimulus/procedure (e.g., Velton mood statements) or reporting a manipulation check. In addition, studies that used erotic and/or sexually arousing stimuli (e.g., pornography) were excluded because other biological systems could have been engaged as well.

3. However, the healthy control groups used for comparison purposes in these studies were included.

4. Two hundred forty-one studies initially qualified for inclusion. However, 29 studies were subsequently excluded because they did not report data from independent samples. When several reports were published that used data from the same original investigation, the article that reported the greatest number of participants or allowed the most accurate calculation of the effect size was included in the meta-analysis. In addition, 4 studies were excluded because they did not provide the necessary information to calculate an effect size. These studies that initially qualified but were subsequently excluded are marked with double asterisks in the References.
1. The time of day was kept constant for all participants (all sessions conducted in the morning or all in the afternoon; $\kappa = 1.00$).

2. A screening procedure was utilized to exclude participants with psychological or physical disorders that could affect the neuroendocrine system (e.g., depression, neuroendocrine disorders, autoimmune disease, cardiovascular disease; $\kappa = .93$).

Other methodological issues may also be important (e.g., behavioral restrictions before the experimental session), but details were often not included in the methods section of the reports and therefore could not be accurately evaluated.

**Characteristics of the Stressor**

Type of stressor task ($\kappa = 1.00$). To explore potential differences in cortisol responses by type of stressor, we classified the stressor task used in each study into one of five categories. In public speaking/verbal interaction tasks, participants were given instructions to verbally interact with an experimenter, confederate, or another participant. This included public speaking tasks, in which participants prepared and delivered a speech on an assigned topic; interviews, in which participants discussed a personal topic such as a negative life experience or an aspect of their personality; and marital conflict interactions, in which couples discussed a problem in their relationship. Cognitive tasks included the Stroop task, or color–word interference task; mental arithmetic tasks; vigilance–reaction time tasks; tests of perceptual skill (e.g., mirror tracing); and other analytical tasks (e.g., anagrams, puzzles). In public speaking/cognitive task combinations, participants delivered a speech and completed a cognitive task consecutively within the same experimental session. Emotion induction procedures included the presentation of emotion-eliciting material designed to automatically elicit a negative affective state (e.g., film), as well as free or guided mental generation of emotional states, in which participants recalled a situation in which they felt a specific affective state, acted out an emotional scenario, or experienced the mood described by a series of statements. In noise exposure tasks, participants experienced either intermittent or continuous loud noise without an accompanying cognitive task.

Duration of the stressor task ($r = .97$). Because the dose of the stressor could affect cortisol responses, the total length (in minutes) of the stressor task was coded and treated as a continuous variable.

Social-evaluative threat ($\kappa = .95$ for composite; reliabilities range from .94 to 1.00 for components). Social-evaluative threat occurs when an aspect of the self (e.g., trait, ability) is or could be negatively judged by others. Elements of the experimental protocol that would induce social-evaluative threat included (a) permanent recording of the performance (videotape or audiotape created the potential for subsequent evaluation), (b) presence of an evaluative audience during the task (at least one other person present besides the experimenter), and (c) presence of a negative social comparison (the real or potential out-performance by a confederate or other participant). These characteristics were coded for each study, creating a dichotomous variable (no social evaluation/social evaluation) and a continuous score, ranging from 0 (no evaluative aspects present) to 3 (three evaluative aspects present).

Uncontrollability ($\kappa = .96$ for composite; reliabilities range from .94 to 1.00 for components). Stressors were defined as uncontrollable when behavioral responses could not appreciably affect outcomes and it would have been reasonable for participants to expect that outcomes were not contingent on their behavior (e.g., it would have been quite difficult for participants to avoid negative consequences, terminate an aversive experience, or succeed despite their best efforts; Henry & Grim, 1990; Weiner, 1992). Specific elements that would inform participants they were failing or could not avoid negative consequences include (a) manipulation of task difficulty (e.g., performing under time constraints, completing impossible tasks); (b) false feedback of poor performance; (c) receiving criticism or harassment from the experimenter during the task regarding their speed, effort expended, overall ability, and so on, that did not include direct feedback about the accuracy of the performance (e.g., right or wrong); and (d) the presence of continuous or intermittent loud noise, auditory distraction, or other emotionally distressing stimuli without the possibility of a behavioral response. These characteristics were coded for each study, creating a dichotomous variable (controllable/uncontrollable) and a continuous score, with a range from 0 (no uncontrollable aspects present) to 4 (four uncontrollable aspects present).

**Motivated performance tasks ($\kappa = 1.00$).** Studies were categorized as passive tasks or motivated performance tasks (Blascovich & Tomaka, 1996). Motivated performance tasks are active performance situations that require or demand immediate overt or cognitive responses on the part of the participant. These tasks are therefore goal relevant and have the potential for evaluation along a self-relevant domain (Blascovich & Mendes, 2000). Examples of motivated performance tasks include completing mental arithmetic problems or delivering a speech. Passive tasks, which do not require instrumental cognitive responses, include watching a film or being exposed to noise (Blascovich & Mendes, 2000).

**Extrapolating and Calculating Effect Sizes**

The standardized mean-change statistic, $d$, was used as the estimate of effect size, which is appropriate for repeated measures effect size estimates (B. J. Becker, 1988; Dunlap, Cortina, Vaslow, & Burke, 1996; Morris, 2000). This statistic can be interpreted as the magnitude of the difference between pre- and poststressor cortisol values in standard deviation units. Effect sizes were calculated with B. J. Becker’s (1988) formula,

$$d = \frac{(M_{poststressor} - M_{prestressor})}{SD_{prestressor}},$$

and corrected for small-sample bias (Hedges & Olkin, 1985). The cortisol measure obtained immediately before the stressor onset was used as the prestressor value, and effect sizes were calculated such that a positive value represents an increase in cortisol levels from pre- to poststressor. Separate $d$s were calculated for each cortisol assessment during and after the stressor (up to 60 min from the end of the stressor). J. Cohen (1988) classifies an effect size of 0.20 as small, 0.50 as moderate, and 0.80 as large.

When possible, effect sizes were calculated from the means and standard deviations reported in the article. When this information was not provided, the effect size was computed from inferential statistics, according to standard meta-analytic procedures (Hedges & Olkin, 1985; Rosenthal, 1991). When the appropriate means and standard deviations or test statistics were not provided, we first attempted to contact the author (or authors) to obtain this information. When the authors could not be contacted or could not provide the requested information ($k = 7$), we proceeded in the most conservative fashion. If “significant changes” were reported and no further information was given, the significance level was assumed to be .05, and this $p$ value was used to calculate the effect size. If “no significant changes” were reported without additional information, the effect size was coded as 0.00 (Rosenthal, 1991). These inferences were made for only three studies (and omitting these studies does not alter the results of the analyses). In four instances, it was impossible to determine whether the stressor evoked significant cortisol changes (e.g., regressions predicting cortisol levels from personality characteristics were reported); these studies were excluded from the research synthesis.
In some cases, only a select group of the study participants was included in the effect size estimates. If a study compared differences in cortisol responses between participants who were recruited on the basis of a psychological or physical disorder or a stressful experience and healthy control participants, only effect sizes for the healthy control group were calculated and used in the analyses. Conditions in which participants received social support (from a friend, confederate, or the experimenter) were not included. In other instances, separate sets of effect size estimates were calculated for a single study. If a between-subjects experiment had several conditions that varied by type of stressor task or characteristics of the stressor (e.g., harassment vs. no harassment), each condition was coded as a separate study, and effect sizes were calculated for each condition.

Within-subjects designs with several stressor tasks or conditions could not be coded separately in the same manner without violating meta-analytic assumptions of independence. Therefore, three rules were used to handle these special cases. First, when the same participants completed the same stress task on multiple days (e.g., habituation paradigms), only effect sizes from the first day of the study were calculated and used in the analyses. Second, when the same participants completed several stressor tasks on the same day that were separated by a rest period greater than 10 min, only effect sizes for the first of these tasks were calculated. This was done in lieu of coding the separate tasks as one longer stressor because it was not clear that intermittent stressors and continuous stressors would result in similar patterns of cortisol changes. (It was not feasible to test differences in intermittent vs. continuous stressors because of the small number of studies that used periods of rest.) Finally, when the same participants completed different stressor tasks on separate days, effect sizes were calculated for only one of the reported sessions. This selection was based on the stressor domain characteristics. After all of the studies were compiled and coded, to maximize the variability, we included the most novel of the stressor tasks (based on the tabulated frequencies of the social-evaluative threat and control dimensions). After all of the studies were compiled and coded, to maximize the variability, we included the most novel of the stressor tasks (based on the tabulated frequencies of the social-evaluative threat and control dimensions).

In total, 599 effect sizes were computed from the 208 studies, with each study contributing, on average, 2.9 effect sizes \(^8\) (range = 1–13, SD = 2.2). The effect sizes represent cortisol responses at different time intervals from stressor onset. The 534 effect sizes calculated from the samples obtained up to 60 min from stressor onset were used in the primary analyses; 505 effect sizes from samples obtained up to 60 min from the termination of the stressor were used in the recovery analyses (440 effect sizes overlap).

In addition, among the subsample of 39 studies that assessed both ACTH and cortisol as indicators of HPA activity, we calculated effect size estimates for changes in ACTH from pre- to poststressor, using the same procedures outlined above. We computed effect sizes for the ACTH assessments obtained 0–60 min from stressor onset; 95 effect sizes were computed, and each study contributed, on average, 2.4 effect sizes (range = 1–5, SD = 1.4).

**Data Analysis**

We adopted the multilevel mixed-model approach outlined in Kalaian and Raudenbush (1996). This model accounts for the hierarchical structure of the data, in which the effect sizes are nested within each study. There are several major advantages to this meta-analytic approach. First, many studies produced multiple effect sizes because of repeated cortisol assessments from participants. This is problematic because effect sizes generated from different assessments within the same study (i.e., from the same participants) would violate statistical assumptions of independence. The hierarchical approach circumvents this problem by (a) accounting for the dependencies in the data; and (b) allowing studies to contribute different numbers of effect sizes, which maximizes the information that can be obtained from each study. Second, multilevel models treat the study as a random, rather than fixed, effect, which permits generalizing the findings to the population rather than restricting them to the particular set of studies included in the research synthesis.

We used Hierarchical Linear Modeling (Raudenbush, Bryk, & Congdon, 2000) and restricted maximum likelihood estimation to analyze the data. Effect sizes were weighted by the sample size of the study \(^6\) (Hedges & Olkin, 1984; Hunter & Schmidt, 1990). Multilevel models in meta-analysis specify two linked equations: an assessment-level component (Level 1), which predicts the effect sizes across the cortisol assessment points, and a study-level component (Level 2), which predicts effect sizes across the studies. First, to estimate the average overall effect size, we tested a baseline model without predictors. The assessment-level component is represented by the Level 1 equation

\[
d_i = \beta_{0i} + R_{0i},
\]

and the study-level component is represented by the Level 2 equation

\[
\beta_{0i} = \gamma_{00} + U_{0i},
\]

where \(d_i\) refers to the effect size for assessment \(i\) from study \(j\), \(\beta_{0i}\) refers to the intercept (average effect size for an average assessment), and \(R_{0i}\) refers to the Level 1 random residual (error). The Level 2 equation represents \(\beta_{0i}\) as a function of the average effect size across all studies (\(\gamma_{00}\)) and the Level 2 random residual (\(U_{0i}\)), or error. The variance of \(U_{0j}\) is a measure of the variability in the effect sizes; a significant variance component indicates that there is a substantial amount of variance unexplained by the model and warrants a search for predictor variables. This test is conceptually similar to the homogeneity statistic (Q) used in other meta-analytic procedures. In this model, the primary coefficients of interest were \(\gamma_{00}\)—the average overall effect size—and \(U_{0i}\)—the variance of the effect size estimate.

In the second phase of analysis, this basic model was then expanded to include predictor variables. There were two different types of predictors: Timing of cortisol assessment was a Level 1 (assessment-level) predictor, and the other coded variables (e.g., time of day the study was conducted, type of stressor task used) were Level 2 (study-level) predictors. Each of the

\(^6\) In cases in which all participants were recruited in the same manner but were then split into separate categories based on gender (e.g., Kirschbaum, Wust, & Hellhammer, 1992), median splits on a questionnaire (e.g., high chronic stress vs. low chronic stress; Benshop et al., 1994), or top and bottom quartiles from a preliminary screen (e.g., high and low heart rate reactors; Sgouros-Emch et al., 1994; high and low hostile men: Suarez, Kuhn, Schanberg, Williams, & Zimmerman, 1998), the groups were combined and included in the meta-analysis.

\(^7\) Seventy-five studies contributed 1 effect size, 42 studies contributed 2 effect sizes, 29 studies contributed 3 effect sizes, 17 studies contributed 4 effect sizes, 19 contributed 5 effect sizes, 13 contributed 6 effect sizes, 7 contributed 7 effect sizes, 1 contributed 8 effect sizes, 2 contributed 9 effect sizes, 2 contributed 11 effect sizes, and 1 contributed 13 effect sizes.

\(^8\) The most common solution to this problem is to aggregate effect sizes within each study across the different outcome variables (e.g., averaging across all cortisol assessments within each study). However, this technique is not appropriate when the relationships between the different outcomes and dependent variable are expected to vary. Cortisol assessments at certain time points from stressor onset could be associated with greater effect sizes, because timing of assessment could play a key role in capturing cortisol responses (Kirschbaum & Hellhammer, 1994). Aggregating effect sizes was therefore not appropriate for this meta-analysis.

This was done because effect size estimates from larger samples are more precise than smaller samples (Hedges & Olkin, 1985). There were four studies with outlying sample size values (i.e., \(z\)-scores greater than 3.29; Tabachnik & Fidell, 1996); these values were brought to 3 standard deviations from the mean before weighting, which substantially reduced the skewness and kurtosis of the sample sizes. Although weighted effect sizes are reported, analyses were also conducted with unweighted effect sizes; this produced virtually equivalent results.
the Level 2 predictors was tested in the same manner; the time of day in which the study was conducted is used as a representative example below. Equations 1 and 2 were expanded to include the predictor variables; the assessment-level (Level 1) component is represented by the equation

\[ d_{ij} = \beta_{0j} + \beta_{1j}(\text{Assessment Timing}) + R_{ij}, \]  

(3)

and the study-level (Level 2) component is represented by the equations

\[ \beta_{0j} = \gamma_{00} + \gamma_{01}(\text{Time of Day}) + U_{0j}, \]  

(4)

and

\[ \beta_{1j} = \gamma_{10} + U_{1j}, \]  

(5)

Terms in the assessment-level equation (3) have corresponding study-level equations (4 and 5). The intercept, \( \beta_{0j} \), is specified by Equation 4, where \( \gamma_{00} \) is the study-level intercept, \( \gamma_{01} \) is an unstandardized regression coefficient relating the effect sizes to time of day the study was conducted, and \( U_{0j} \) is the study-level random residual. An unstandardized regression coefficient, \( \beta_{1j} \), relating the effect sizes to the time of cortisol assessment, is specified by Equation 5, where \( \gamma_{10} \) represents the average effect of assessment timing across all the studies and \( U_{1j} \) represents the residual between-groups random component. In the examination of predictor variables, the primary regression coefficients of interest were \( \gamma_{10} \) (effect of timing of cortisol assessment) and \( \gamma_{11} \) (effect of study-level predictor, e.g., time of day).

When reporting average effect sizes (either for the synthesis as a whole or a subsample of studies with specific characteristics), we present the average \( d \), 95% confidence interval (CI), and \( p \) value. When reporting specific study-level predictors or contrasts between studies with different characteristics, we present \( \gamma_{0j} \), the \( t \) test, and \( p \) value.

**Results**

**Descriptives**

In total, 6,153 individuals participated in the 208 studies included in the meta-analysis. On average, 57% of the participants in each study were male (range = 0%–100%, \( SD = 37 \)), and the mean age of the participants was 30.8 years (range = 18.4–72.3, \( SD = 12.0 \)). A slight majority of the studies (115; 55%) assessed cortisol through plasma (rather than salivary) measures, with the average cortisol assessment occurring 29.9 min (\( SD = 16.2 \)) from stressor onset. In terms of the type of stressor task used, 101 studies (48%) used cognitive tasks, 37 (18%) used public speaking/verbal interaction tasks, 48 (23%) used public speaking/cognitive task combinations, 16 (8%) used emotion induction techniques, and 6 (3%) used noise exposure.

<table>
<thead>
<tr>
<th>Type of stressor task</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>No. with social-evaluative threat</th>
<th>No. with uncontrollability</th>
<th>Length of stressor task (min)</th>
<th>Time of day conducted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive tasks</td>
<td>101</td>
<td>2,480</td>
<td>14 (14%)</td>
<td>76 (75%)</td>
<td>22.3 (14.6)</td>
<td>41 AM, 41 PM, 19 NA</td>
</tr>
<tr>
<td>Public speaking/verbal interaction</td>
<td>37</td>
<td>1,629</td>
<td>34 (92%)</td>
<td>1 (3%)</td>
<td>18.8 (11.7)</td>
<td>15 AM, 17 PM, 5 NA</td>
</tr>
<tr>
<td>Public speaking/cognitive combination</td>
<td>48</td>
<td>1,553</td>
<td>45 (94%)</td>
<td>43 (90%)</td>
<td>20.2 (10.8)</td>
<td>19 AM, 24 PM, 5 NA</td>
</tr>
<tr>
<td>Noise exposure</td>
<td>6</td>
<td>109</td>
<td>0 (0%)</td>
<td>5 (83%)</td>
<td>35.0 (18.4)</td>
<td>2 AM, 4 PM</td>
</tr>
<tr>
<td>Emotion induction</td>
<td>16</td>
<td>382</td>
<td>1 (6%)</td>
<td>8 (50%)</td>
<td>17.1 (12.9)</td>
<td>6 AM, 8 PM, 2 NA</td>
</tr>
<tr>
<td>Total</td>
<td>208</td>
<td>6,153</td>
<td>94 (45%)</td>
<td>133 (64%)</td>
<td>21.2 (13.5)</td>
<td>83 AM, 94 PM, 31 NA</td>
</tr>
</tbody>
</table>

**Note.** For length of stressor tasks, standard deviations are in parentheses. AM = morning; PM = afternoon; NA = not assessed or not available.

Table 1 displays descriptive statistics for each stressor task category and the sample as a whole, including the number of participants, characteristics of the stressor task (length of stressor, social-evaluative threat, and uncontrollability), and the frequencies for the time of day the study was conducted. Table 2 displays the average effect size and primary coded dimensions for each study.

**Overall Effect**

The average effect size across all assessments from all studies was 0.31 (\( SEM = 0.044 \)), which is significantly different from zero, \( n(207) = 7.10, p < .01; 95\% CI = 0.22, 0.40 \). This demonstrates that acute laboratory-induced psychological stressors elicit a significant cortisol response, on average increasing cortisol levels 0.31 standard deviations above prestressor baseline values. The variance component was also significant (0.30), \( \chi^2(207, N = 208) = 1,330.38, p < .01 \), which suggests that there is variance in cortisol changes that could be predicted by other factors.

**Predicting the Effect Sizes: Methodological Factors**

We tested two categories of predictor variables: methodological factors (e.g., timing of cortisol assessment, time of day the study was conducted) and the characteristics of the stressor task (e.g., type of task, uncontrollability). We first tested the methodological predictors, because these design features could potentially mask relationships between cortisol responses and the characteristics of the stressor tasks. We performed regression analyses to test whether methodological factors predicted the effect sizes (the \( \gamma \) reported can be interpreted as unstandardized regression coefficients). After establishing a relationship between any methodological factors and cortisol responses, we then tested, using log-likelihood ratios, whether including specific sets of methodological factors in a model predicting effect sizes improved the overall model fit (e.g., explained more variance). These procedures allowed us to identify a set of methodological factors that best explained the variability in effect sizes and could then be used as controls in subsequent analyses.

**Individual Predictors**

Table 3 provides a summary of the regression coefficients and significance tests for the methodological predictors. We first tested (text continues on page 368)
Table 2

Coded Dimensions and Average Effect Sizes for the Studies Included in the Research Synthesis

<table>
<thead>
<tr>
<th>Study and task type</th>
<th>d</th>
<th>No. of d</th>
<th>N</th>
<th>Time of day</th>
<th>Category</th>
<th>SET</th>
<th>UC</th>
<th>Type</th>
<th>Length (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abplanalp et al. (1977)</td>
<td>1.20</td>
<td>1</td>
<td>21</td>
<td>—</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Ackerman et al. (1996)</td>
<td>−0.16</td>
<td>3</td>
<td>25</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>5</td>
</tr>
<tr>
<td>Adlercreutz et al. (1982)</td>
<td>1.54</td>
<td>2</td>
<td>6</td>
<td>AM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>al’Absi et al. (1994)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermittent condition</td>
<td>0.06</td>
<td>1</td>
<td>20</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>35</td>
</tr>
<tr>
<td>al’Absi et al. (1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental arithmetic</td>
<td>1.31</td>
<td>2</td>
<td>23</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>24</td>
</tr>
<tr>
<td>Public speaking</td>
<td>0.68</td>
<td>2</td>
<td>23</td>
<td>—</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>24</td>
</tr>
<tr>
<td>al’Absi et al. (2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public speaking</td>
<td>1.53</td>
<td>1</td>
<td>46</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>24</td>
</tr>
<tr>
<td>Altemus, Rao, et al. (2001)</td>
<td>0.46</td>
<td>2</td>
<td>25</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Altemus, Redwine, et al. (2001)</td>
<td>0.93</td>
<td>5</td>
<td>14</td>
<td>—</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Andren et al. (1982)</td>
<td>−0.16</td>
<td>1</td>
<td>15</td>
<td>PM</td>
<td>NO</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>20</td>
</tr>
<tr>
<td>Arguelles et al. (1962)</td>
<td>1.09</td>
<td>1</td>
<td>6</td>
<td>AM</td>
<td>NO</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>60</td>
</tr>
<tr>
<td>Arnetz et al. (1985)</td>
<td>0.27</td>
<td>3</td>
<td>10</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>50</td>
</tr>
<tr>
<td>Arnetz et al. (1986b)</td>
<td>−0.02</td>
<td>3</td>
<td>22</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>50</td>
</tr>
<tr>
<td>Bartholomew (1997)</td>
<td>−0.49</td>
<td>1</td>
<td>20</td>
<td>—</td>
<td>EI</td>
<td>No</td>
<td>No</td>
<td>PA</td>
<td>20</td>
</tr>
<tr>
<td>L. C. Becker et al. (1996)</td>
<td>0.26</td>
<td>1</td>
<td>29</td>
<td>AM</td>
<td>PS/CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Benshop et al. (1994)</td>
<td>−0.44</td>
<td>2</td>
<td>27</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Altemus, Rao, et al. (2001)</td>
<td>0.46</td>
<td>2</td>
<td>25</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Altemus, Redwine, et al. (2001)</td>
<td>0.93</td>
<td>5</td>
<td>14</td>
<td>—</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Andren et al. (1982)</td>
<td>−0.16</td>
<td>1</td>
<td>15</td>
<td>PM</td>
<td>NO</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>20</td>
</tr>
<tr>
<td>Arguelles et al. (1962)</td>
<td>1.09</td>
<td>1</td>
<td>6</td>
<td>AM</td>
<td>NO</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>60</td>
</tr>
<tr>
<td>Arnetz et al. (1985)</td>
<td>0.27</td>
<td>3</td>
<td>10</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>50</td>
</tr>
<tr>
<td>Arnetz et al. (1986b)</td>
<td>−0.02</td>
<td>3</td>
<td>22</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>50</td>
</tr>
<tr>
<td>Bartholomew (1997)</td>
<td>−0.49</td>
<td>1</td>
<td>20</td>
<td>—</td>
<td>EI</td>
<td>No</td>
<td>No</td>
<td>PA</td>
<td>20</td>
</tr>
<tr>
<td>L. C. Becker et al. (1996)</td>
<td>0.26</td>
<td>1</td>
<td>29</td>
<td>AM</td>
<td>PS/CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Benshop et al. (1994)</td>
<td>−0.44</td>
<td>2</td>
<td>27</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>30</td>
</tr>
</tbody>
</table>

364 DICKERSON AND KEMENY
<table>
<thead>
<tr>
<th>Study and task type</th>
<th>No. of subjects</th>
<th>N</th>
<th>Time of day</th>
<th>Category</th>
<th>SET</th>
<th>UC</th>
<th>Type</th>
<th>Length (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earle et al. (1999)</td>
<td>21</td>
<td>30</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>No harassment</td>
<td>-26</td>
<td>30</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>Eilenbogen et al. (2002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative stressor</td>
<td>0.00</td>
<td>47</td>
<td>PM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>35</td>
</tr>
<tr>
<td>Positive stressor</td>
<td>-0.08</td>
<td>45</td>
<td>PM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>35</td>
</tr>
<tr>
<td>Neutral stressor</td>
<td>-0.01</td>
<td>43</td>
<td>PM</td>
<td>CT</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>35</td>
</tr>
<tr>
<td>Epel et al. (1999)</td>
<td>-0.02</td>
<td>27</td>
<td>AM</td>
<td>PS/CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>60</td>
</tr>
<tr>
<td>No harassment</td>
<td>0.26</td>
<td>4</td>
<td>AM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>Ellenbogen et al. (2002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative stressor</td>
<td>0.00</td>
<td>52</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>45</td>
</tr>
<tr>
<td>Positive stressor</td>
<td>0.08</td>
<td>24</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Neutral stressor</td>
<td>-0.88</td>
<td>160</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Epel et al. (2000)</td>
<td>0.15</td>
<td>5</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>45</td>
</tr>
<tr>
<td>Fehm-Wolfsdorf et al. (1993)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fehm-Wolfsdorf et al. (1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibiger et al. (1986)</td>
<td>0.03</td>
<td>8</td>
<td></td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>25</td>
</tr>
<tr>
<td>Fountine (2001)</td>
<td>0.01</td>
<td>27</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Fredrikson &amp; Blumenthal (1992)</td>
<td>0.48</td>
<td>33</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Furlan et al. (2001)</td>
<td>0.23</td>
<td>17</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Futterman et al. (1994)</td>
<td>-0.47</td>
<td>14</td>
<td>AM</td>
<td>EI</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Gallinelli et al. (2000)</td>
<td>0.30</td>
<td>20</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Gerra et al. (1996)</td>
<td>0.00</td>
<td>20</td>
<td>AM</td>
<td>EI</td>
<td>Yes</td>
<td>No</td>
<td>PA</td>
<td>45</td>
</tr>
<tr>
<td>Gerra et al. (1997)</td>
<td>0.37</td>
<td>30</td>
<td>—</td>
<td>CT</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Gerra et al. (1998)</td>
<td>0.22</td>
<td>16</td>
<td>PM</td>
<td>NO</td>
<td>Yes</td>
<td>No</td>
<td>PA</td>
<td>30</td>
</tr>
<tr>
<td>Gerra et al. (1999)</td>
<td>0.19</td>
<td>14</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>30</td>
</tr>
<tr>
<td>Gerra et al. (2001)</td>
<td>2.33</td>
<td>20</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Gerritsen et al. (1996)</td>
<td>-0.05</td>
<td>35</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>45</td>
</tr>
<tr>
<td>Study 1</td>
<td></td>
<td></td>
<td></td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>50</td>
</tr>
<tr>
<td>Failure condition</td>
<td>-0.10</td>
<td>38</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>50</td>
</tr>
<tr>
<td>Gotthardt et al. (1995)</td>
<td>1.36</td>
<td>40</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>45</td>
</tr>
<tr>
<td>Green (2001)</td>
<td>-0.32</td>
<td>40</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>3</td>
</tr>
<tr>
<td>Griffiths et al. (1997)</td>
<td>0.24</td>
<td>10</td>
<td>AM</td>
<td>PS</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>7</td>
</tr>
<tr>
<td>Halpern et al. (2002)</td>
<td>-0.05</td>
<td>142</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Hawkley et al. (2001)</td>
<td>0.43</td>
<td>64</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>Heesen et al. (2002)</td>
<td>-0.07</td>
<td>15</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>45</td>
</tr>
<tr>
<td>Heim et al. (2000)</td>
<td>0.51</td>
<td>12</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Hellhammer et al. (1997)</td>
<td>0.52</td>
<td>63</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Hill (2000)</td>
<td>-0.20</td>
<td>128</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Hoehn et al. (1997)</td>
<td>-0.50</td>
<td>10</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>24</td>
</tr>
<tr>
<td>Hoehn-Saric et al. (1991)</td>
<td>-0.04</td>
<td>14</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Holt et al. (1984)</td>
<td>2.36</td>
<td>5</td>
<td>AM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Hollenberg et al. (1981)</td>
<td>-0.31</td>
<td>23</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>26</td>
</tr>
<tr>
<td>Horan (2002)</td>
<td>0.19</td>
<td>15</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>Hubert &amp; de Jong-Meyer (1990)</td>
<td>-0.70</td>
<td>12</td>
<td>PM</td>
<td>EI</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>9</td>
</tr>
<tr>
<td>Hubert &amp; de Jong-Meyer (1991)</td>
<td>-1.01</td>
<td>20</td>
<td>PM</td>
<td>EI</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>10</td>
</tr>
<tr>
<td>Hucklebridge et al. (2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study 1: Recall</td>
<td>0.00</td>
<td>2</td>
<td>19 AM</td>
<td>EI</td>
<td>No</td>
<td>No</td>
<td>PA</td>
<td>10</td>
</tr>
<tr>
<td>Study 2: Music</td>
<td>0.15</td>
<td>1</td>
<td>35 AM</td>
<td>EI</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>30</td>
</tr>
<tr>
<td>Hyypa et al. (1983)</td>
<td>-0.63</td>
<td>18</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>60</td>
</tr>
<tr>
<td>Jansen et al. (1998)</td>
<td>0.89</td>
<td>10</td>
<td>—</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>25</td>
</tr>
<tr>
<td>Jansen et al. (2000)</td>
<td>0.49</td>
<td>21</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>40</td>
</tr>
<tr>
<td>Jones et al. (1997)</td>
<td>0.35</td>
<td>39</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Jorgensen et al. (1990)</td>
<td>0.21</td>
<td>14</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Jorgensen et al. (1993)</td>
<td>0.00</td>
<td>14</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Kaciuba-Uscilko et al. (1994)</td>
<td>-0.19</td>
<td>20</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Kahn et al. (1992)</td>
<td>1.32</td>
<td>10</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>40</td>
</tr>
<tr>
<td>Kang &amp; Fox (2000)</td>
<td>-0.04</td>
<td>16</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>18</td>
</tr>
<tr>
<td>Kemmer et al. (1986)</td>
<td>0.15</td>
<td>3</td>
<td>9 AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Public speaking</td>
<td>0.03</td>
<td>62</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Kiecolt-Glaser et al. (1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study and task type</td>
<td>d</td>
<td>No. of d</td>
<td>N</td>
<td>Time of day</td>
<td>Category</td>
<td>SET</td>
<td>UC</td>
<td>Type</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----</td>
<td>----------</td>
<td>-----</td>
<td>-------------</td>
<td>----------</td>
<td>-----</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>Kirschbaum, Bartussek, &amp; Strasburger (1992)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study 1</td>
<td>0.85</td>
<td>6</td>
<td>50</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Study 2</td>
<td>1.09</td>
<td>6</td>
<td>37</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Kirschbaum et al. (1994)</td>
<td>0.66</td>
<td>5</td>
<td>20</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Kirschbaum, Klauer, et al. (1995)</td>
<td>1.05</td>
<td>5</td>
<td>23</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Study 1</td>
<td>1.16</td>
<td>1</td>
<td>12</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Study 2</td>
<td>0.83</td>
<td>1</td>
<td>64</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Kirschbaum et al. (1996)</td>
<td>2.30</td>
<td>5</td>
<td>20</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Kirschbaum et al. (1999)</td>
<td>0.95</td>
<td>4</td>
<td>81</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Knight &amp; Rickard (2001)</td>
<td>0.43</td>
<td>1</td>
<td>42</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Korchin &amp; Herz (1960)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scrambled sentences</td>
<td>−0.32</td>
<td>1</td>
<td>10</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Picture description</td>
<td>0.43</td>
<td>1</td>
<td>10</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Kudielka et al. (2000)</td>
<td>1.24</td>
<td>5</td>
<td>27</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Larson et al. (2001)</td>
<td>−0.12</td>
<td>1</td>
<td>55</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Lehmann et al. (1992)</td>
<td>−0.41</td>
<td>3</td>
<td>17</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Leyton et al. (1996)</td>
<td>−0.30</td>
<td>6</td>
<td>8</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Linden &amp; Long (1987)</td>
<td>0.09</td>
<td>3</td>
<td>57</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Lovatto et al. (1985)</td>
<td>1.32</td>
<td>1</td>
<td>28</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Lovatto et al. (1986b)</td>
<td>0.09</td>
<td>1</td>
<td>43</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Lovatto et al. (2000)</td>
<td>0.75</td>
<td>2</td>
<td>10</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Luecken (1998)</td>
<td>−0.21</td>
<td>2</td>
<td>31</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Lupien et al. (1997)</td>
<td>−0.15</td>
<td>3</td>
<td>14</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Malarkey et al. (1994)</td>
<td>−0.12</td>
<td>4</td>
<td>180</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Manuck et al. (1991)</td>
<td>0.13</td>
<td>1</td>
<td>20</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Marinari et al. (1976)</td>
<td>0.51</td>
<td>1</td>
<td>60</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Mathe &amp; Knapp (1971)</td>
<td>1.00</td>
<td>1</td>
<td>6</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Matthews et al. (2001)</td>
<td>−0.01</td>
<td>2</td>
<td>62</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>McCleary et al. (2000)</td>
<td>−0.12</td>
<td>4</td>
<td>30</td>
<td>PM</td>
<td>EI</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
</tr>
<tr>
<td>Miki et al. (1998)</td>
<td>2.46</td>
<td>3</td>
<td>8</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Noise</td>
<td>−0.16</td>
<td>2</td>
<td>82</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Miller et al. (1999)</td>
<td>0.13</td>
<td>6</td>
<td>9</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Miyabo et al. (1976)</td>
<td>0.04</td>
<td>7</td>
<td>20</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Miyabo et al. (1979)</td>
<td>0.50</td>
<td>4</td>
<td>5</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Moyer et al. (1994)</td>
<td>−0.13</td>
<td>3</td>
<td>41</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Moyna et al. (1999)</td>
<td>0.07</td>
<td>2</td>
<td>45</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Nejkov (2002)</td>
<td>0.30</td>
<td>2</td>
<td>44</td>
<td>PM</td>
<td>EI</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
</tr>
<tr>
<td>Neumann et al. (1992)</td>
<td>−0.37</td>
<td>3</td>
<td>15</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Neumann et al. (1994)</td>
<td>−0.30</td>
<td>2</td>
<td>20</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Nicolai et al. (1997)</td>
<td>0.13</td>
<td>3</td>
<td>54</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Odink et al. (1987)</td>
<td>0.72</td>
<td>1</td>
<td>32</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Pasquali et al. (1996)</td>
<td>0.00</td>
<td>4</td>
<td>6</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Peters et al. (1998)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High effort, control</td>
<td>−0.29</td>
<td>3</td>
<td>21</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>High effort, no control</td>
<td>−0.19</td>
<td>3</td>
<td>22</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Low effort, control</td>
<td>−0.37</td>
<td>3</td>
<td>20</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Low effort, no control</td>
<td>−0.15</td>
<td>3</td>
<td>22</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Pettingale et al. (1989)</td>
<td>−0.26</td>
<td>1</td>
<td>9</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Pike et al. (1997)</td>
<td>1.02</td>
<td>2</td>
<td>23</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Pirke et al. (1992)</td>
<td>0.87</td>
<td>1</td>
<td>20</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Pruessner et al. (1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Failure</td>
<td>0.80</td>
<td>4</td>
<td>26</td>
<td>PM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Success</td>
<td>1.01</td>
<td>4</td>
<td>26</td>
<td>PM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Raab (1968)</td>
<td>0.09</td>
<td>2</td>
<td>47</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Ravindran et al. (1996)</td>
<td>−0.07</td>
<td>1</td>
<td>23</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Riet et al. (1998)</td>
<td>−0.19</td>
<td>1</td>
<td>19</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Rohleder et al. (2001)</td>
<td>1.12</td>
<td>5</td>
<td>45</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Rohrhann et al. (1999)</td>
<td>−0.06</td>
<td>1</td>
<td>20</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
</tr>
<tr>
<td>Roy et al. (1998)</td>
<td>−0.30</td>
<td>2</td>
<td>93</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
</tr>
<tr>
<td>Sachs et al. (1993)</td>
<td>−0.33</td>
<td>1</td>
<td>18</td>
<td>AM</td>
<td>PS</td>
<td>No</td>
<td>No</td>
<td>MP</td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th>Study and task type</th>
<th>d</th>
<th>No. of ds</th>
<th>N</th>
<th>Time of day</th>
<th>Category</th>
<th>SET</th>
<th>UC</th>
<th>Type</th>
<th>Length (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sauro (2002)</td>
<td>0.27</td>
<td>3</td>
<td>34</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Low threat</td>
<td>-0.16</td>
<td>3</td>
<td>34</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Scarpa et al. (2000)</td>
<td>-0.39</td>
<td>1</td>
<td>54</td>
<td>PM</td>
<td>NO</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>55</td>
</tr>
<tr>
<td>Schmidt-Ott et al. (1998)</td>
<td>0.39</td>
<td>1</td>
<td>7</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Schmidt-Reinwald et al. (1999)</td>
<td>1.61</td>
<td>5</td>
<td>22</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>13</td>
</tr>
<tr>
<td>Schommer et al. (1998)</td>
<td>0.77</td>
<td>5</td>
<td>81</td>
<td>—</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>13</td>
</tr>
<tr>
<td>Seeman, Berkman, et al. (1995)</td>
<td>1.70</td>
<td>3</td>
<td>16</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>40</td>
</tr>
<tr>
<td>Seeman et al. (2001)</td>
<td>0.45</td>
<td>4</td>
<td>33</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Sepon (1995)</td>
<td>-0.34</td>
<td>3</td>
<td>10</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Sgoutas (1992)</td>
<td>-0.28</td>
<td>3</td>
<td>36</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>8</td>
</tr>
<tr>
<td>Sgoutas-Emch et al. (1994)</td>
<td>0.24</td>
<td>1</td>
<td>22</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>Sharples &amp; McLean (1992)</td>
<td>0.16</td>
<td>5</td>
<td>19</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>8</td>
</tr>
<tr>
<td>Sjogren et al. (1999)</td>
<td>1.78</td>
<td>5</td>
<td>27</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Sinyor et al. (2000)</td>
<td>-0.03</td>
<td>7</td>
<td>30</td>
<td>AM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>17</td>
</tr>
<tr>
<td>Skovnik et al. (2000)</td>
<td>-0.18</td>
<td>2</td>
<td>20</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>55</td>
</tr>
<tr>
<td>Sothmann et al. (1998)</td>
<td>-0.03</td>
<td>5</td>
<td>17</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Steptoe et al. (1993)</td>
<td>0.11</td>
<td>2</td>
<td>20</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>5</td>
</tr>
<tr>
<td>No time pressure</td>
<td>0.01</td>
<td>2</td>
<td>20</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>5</td>
</tr>
<tr>
<td>Steptoe et al. (1996)</td>
<td>-0.06</td>
<td>4</td>
<td>66</td>
<td>PM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>27</td>
</tr>
<tr>
<td>Steptoe et al. (2001)</td>
<td>-0.65</td>
<td>3</td>
<td>13</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Stone et al. (1999)</td>
<td>-0.41</td>
<td>1</td>
<td>15</td>
<td>PM</td>
<td>PS</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Stoney et al. (1999)</td>
<td>0.14</td>
<td>2</td>
<td>100</td>
<td>AM</td>
<td>CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>3</td>
</tr>
<tr>
<td>Strood et al. (2000)</td>
<td>0.47</td>
<td>4</td>
<td>22</td>
<td>PM</td>
<td>PS</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>35</td>
</tr>
<tr>
<td>Stojd et al. (2002)</td>
<td>0.32</td>
<td>3</td>
<td>25</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>45</td>
</tr>
<tr>
<td>Suarez et al. (1998)</td>
<td>-0.01</td>
<td>2</td>
<td>23</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Suarez &amp; Harralson (1999)</td>
<td>1.79</td>
<td>1</td>
<td>36</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>11</td>
</tr>
<tr>
<td>Testa et al. (1994)</td>
<td>1.58</td>
<td>1</td>
<td>8</td>
<td>AM</td>
<td>NO</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>15</td>
</tr>
<tr>
<td>Thorsteinsson et al. (1998)</td>
<td>-0.63</td>
<td>2</td>
<td>20</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>5</td>
</tr>
<tr>
<td>Trestman et al. (1991)</td>
<td>1.10</td>
<td>4</td>
<td>12</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>25</td>
</tr>
<tr>
<td>Tsuda et al. (1996)</td>
<td>-0.25</td>
<td>2</td>
<td>20</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>Uchino et al. (1995)</td>
<td>0.05</td>
<td>1</td>
<td>23</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>12</td>
</tr>
<tr>
<td>Usuiyama et al. (1991)</td>
<td>0.25</td>
<td>1</td>
<td>26</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>3</td>
</tr>
<tr>
<td>VanderArk et al. (1993)</td>
<td>0.00</td>
<td>1</td>
<td>60</td>
<td>AM</td>
<td>EI</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>5</td>
</tr>
<tr>
<td>van der Pompe et al. (1996)</td>
<td>-0.32</td>
<td>3</td>
<td>15</td>
<td>AM</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>8</td>
</tr>
<tr>
<td>van Eck et al. (1996)</td>
<td>0.76</td>
<td>4</td>
<td>87</td>
<td>—</td>
<td>PS</td>
<td>Yes</td>
<td>No</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>van Hon et al. (2000)</td>
<td>0.25</td>
<td>1</td>
<td>39</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>10</td>
</tr>
<tr>
<td>Walker (1996)</td>
<td>-0.14</td>
<td>1</td>
<td>10</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Williams et al. (1982)</td>
<td>0.33</td>
<td>1</td>
<td>21</td>
<td>—</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>20</td>
</tr>
<tr>
<td>Williams et al. (1991)</td>
<td>0.28</td>
<td>2</td>
<td>28</td>
<td>AM</td>
<td>CT</td>
<td>No</td>
<td>No</td>
<td>MP</td>
<td>18</td>
</tr>
<tr>
<td>Wittersheim et al. (1985)</td>
<td>0.66</td>
<td>6</td>
<td>10</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Mental arithmetic</td>
<td>0.62</td>
<td>6</td>
<td>10</td>
<td>PM</td>
<td>CT</td>
<td>No</td>
<td>Yes</td>
<td>MP</td>
<td>30</td>
</tr>
<tr>
<td>Multiple choice task</td>
<td>1.96</td>
<td>1</td>
<td>22</td>
<td>AM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Wolf et al. (2001)</td>
<td>0.72</td>
<td>10</td>
<td>10</td>
<td>PM</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Young et al. (2000)</td>
<td>0.72</td>
<td>10</td>
<td>47</td>
<td>—</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Young &amp; Nolen-Hoekelessa (2001)</td>
<td>0.72</td>
<td>10</td>
<td>47</td>
<td>—</td>
<td>PS/CT</td>
<td>Yes</td>
<td>Yes</td>
<td>MP</td>
<td>15</td>
</tr>
<tr>
<td>Zakowski et al. (1992)</td>
<td>-0.44</td>
<td>1</td>
<td>20</td>
<td>AM</td>
<td>EI</td>
<td>No</td>
<td>Yes</td>
<td>PA</td>
<td>30</td>
</tr>
</tbody>
</table>

Note. d represents the average effect size from cortisol assessments 1–60 min from stressor onset; number of ds is the number of effect sizes contributing to this average. Dashes indicate not assessed or not available. SET = social-evaluative threat; UC = uncontrollability; AM = morning; PM = afternoon; PS = public speaking/verbal interaction task; CT = cognitive task; PS/CT = public speaking/cognitive combination task; NO = noise exposure; EI = emotion induction; MP = motivated performance task; PA = passive task.
whether the timing of cortisol assessment, measured as minutes from stressor onset, predicted the within-study differences in effect sizes (variability across cortisol assessments). Figure 1 displays the average effect size for the cortisol assessments obtained during each 10-min epoch from stressor onset. The linear effect of time was nonsignificant. However, there was a significant quadratic effect. Focused contrasts revealed that the cortisol assessments obtained 21–40 min from stressor onset were significantly greater than those from other assessment time points (γ_{10} = 0.10, SEM = 0.022), t(207) = 4.62, p < .01. This indicates that the peak cortisol response occurs 21–40 min from onset of acute psychological stressors.  

Because cortisol displays diurnal variation, it could be more difficult to detect a cortisol response in the morning (when levels naturally decrease) than in the afternoon (when levels are relatively stable). Consistent with this reasoning, time of day significantly predicted effect sizes, accounting for 9% of the between-studies variance. The studies conducted in the morning had an average effect size of 0.14 (CI = 0.03, 0.25, p < .05). Those conducted in the afternoon were associated with greater cortisol changes; the average effect size was 0.46 (CI = 0.31, 0.61, p < .01). This difference in effect size corresponds to obtaining a small cortisol response in the morning and a moderate cortisol response in the afternoon (J. Cohen, 1988), and indicates that time of day is an important methodological factor to consider when designing and/or interpreting the results of acute psychological stressor studies.

We also examined whether participant characteristics (the mean age and gender composition of the study participants) or the method of cortisol assessment (plasma vs. saliva) were associated with cortisol changes. However, none of these characteristics significantly predicted effect sizes (ps > .20), suggesting that these factors do not explain the variability in the cortisol response between studies.

Finally, we tested whether several other methodological features were associated with cortisol changes. Conducting all sessions at the same time of day or excluding participants with conditions that could affect the neuroendocrine system did not predict cortisol responses (ps > .20). This indicates that these methodological features of the study, designed to reduce error variance, did not differ from the others in the size of the overall cortisol response.

### Model Comparison

Individual tests of the methodological characteristics revealed two significant predictors: timing of assessment from stressor onset and time of day the study was conducted. We compared a model that included these predictors with the baseline model (which predicted effect sizes only from the intercept, or average effect size) to determine whether including these methodological characteristics improved the model fit to the data. Consistent with the regression analyses, including both time of day and timing of assessment from stressor onset as predictors significantly improved the fit of the model, as shown in the second column of Table 4. Furthermore, omitting either the time of day or timing of assessment predictors led to a significant degradation in model fit (p < .01). Including additional methodological predictors did not improve the fit (p > .20). This indicates that the model including the timing of assessment and time of day predictors obtains the best, most parsimonious fit to the data. Consequently, these two factors were retained as methodological control variables in the subsequent analyses.

### Predicting the Effect Sizes: Characteristics of the Stressor Tasks

We conducted regression analyses to test whether the characteristics of the stressor tasks (e.g., type of task, uncontrollability) predicted the effect sizes, controlling for the time of day the study was conducted and timing of cortisol assessment from stressor onset. Then, we compared a model that included the methodological and stressor task predictors with one that included only the methodological predictors, which tested whether the addition of stressor task predictors increased the overall fit to the data. Finally, we tested whether the categorical (type of task) or stressor domain (uncontrollability, social-evaluative threat) characteristics were

---

10 To examine whether the time patterns were present because some samples were collected during the stressor, we recalculated the analyses using a reduced set of poststressor samples (n = 440). The same effects were obtained; there was a significant quadratic effect, with a peak 21–40 min from stressor onset.
stronger predictors of the effect sizes through both simultaneous regression and log-likelihood ratios, in which a significant chi-square indicates differences in predictive strength.

Type of Task

Stressors were categorized as cognitive tasks, public speaking/verbal interaction tasks, public speaking/cognitive task combinations, emotion induction procedures, or noise exposure. Figure 2 shows the average effect size for each task category. Significant cortisol responses were elicited by the cognitive tasks \( (d = 0.20, \ CI = 0.09, 0.31, p < .01) \), public speaking/verbal interaction tasks \( (d = 0.39, \ CI = 0.15, 0.63, p < .01) \), and the public speaking/cognitive task combinations \( (d = 0.87, \ CI = 0.68, 1.06, p < .01) \). However, the average effect sizes for the emotion induction \( (d = -0.13) \) and noise exposure \( (d = -0.06) \) categories were not significantly different from zero \( (\ CI = -0.35, 0.09, p > .20; \ CI = -0.49, 0.37, p > .20, \ \text{respectively}) \). Although this indicates that, on average, the emotion induction and noise exposure stressors did not elicit a significant cortisol response, this should be interpreted with caution because of the relatively small numbers of studies that fall in these categories.

We then tested whether certain categories of stressor tasks were associated with greater cortisol changes. When the stressor categories were dummy coded into orthogonal contrasts and simultaneously entered into the regression equation (with the methodological controls), the public speaking/cognitive task combination stressors had a larger average effect size than the other stressor categories \( (\gamma_{01} = 0.039, \ SEM = 0.0081), t(202) = 4.76, p < .01 \), with an effect size over 2 times as large as any other. The orthogonal contrasts revealed no differences in effect sizes between the other categories \( (p > .20) \). These analyses indicate that the public speaking/cognitive task combinations elicited greater cortisol changes than the other types of stressors; however, the stressor category cannot reliably differentiate cortisol responses between other types of tasks.

Length of Stressor

We also tested whether the length of the stressor was associated with cortisol changes. However, stressor length did not significantly predict effect sizes \( (\gamma_{01} = 0.0021, \ SEM = 0.0026), t(205) = 0.79, p > .20 \). This indicates that, overall, the longer stressor tasks were not associated with greater cortisol responses than the shorter tasks. When we controlled for the length of the stressor task, the contrast between public speaking/cognitive tasks and other tasks remained a significant predictor of the effect sizes \( (p < .01) \), suggesting that stressor length does not explain the association between this type of task and cortisol responses.

Stressor Domains

On the basis of the theoretical literature, we hypothesized that contexts characterized by social-evaluative threat (an aspect of the self is or could be negatively judged by others) or uncontrollability (a context of forced failure in which participants could not succeed despite their best efforts) would reliably elicit large cortisol changes.

Social-evaluative threat. Consistent with hypotheses, conditions of social-evaluative threat, in which the performance was captured on permanent record (e.g., videotape), an evaluative audience was present, or a person offering negative social comparison was present, significantly predicted the effect sizes \( (\gamma_{01} = 0.27, \ SEM = 0.075), t(205) = 3.59, p < .01 \). Although both categories significantly increased cortisol levels, tasks with social-evaluative threat elicited a significantly greater cortisol response \( (d = 0.67, \ CI = 0.50, 0.84, p < .01) \) than those without this characteristic \( (d = 0.15, \ CI = 0.02, 0.27, p < .05) \).

Of the studies with social evaluation, some had one form of social evaluation, whereas others had two (e.g., speech was videotaped and performed in front of an audience). We then tested whether the addition of evaluative elements was associated with an increase in the cortisol response. There was a significant difference in effect size between studies with one and two forms of social
Table 4  
Comparison of Models Predicting Effect Sizes

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Log-likelihood</th>
<th>Parameters</th>
<th>Parameters estimated</th>
<th>Log-likelihood ratio</th>
<th>$\chi^2$, $\text{df} \times 10^1$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(no predictors)</td>
<td>$-88.8$</td>
<td>0</td>
<td>$\chi^2(0) = 208$, $p &gt; .01$</td>
<td>$\chi^2(1, N = 208) = 3.42$, $p &lt; .01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(timing of assessment, time of day)</td>
<td>$-88.48$</td>
<td>3</td>
<td>$\chi^2(3, N = 208) = 2.27$, $p &gt; .01$</td>
<td>$\chi^2(1, N = 208) = 3.32$, $p &lt; .01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>(uncontrollability, social-evaluative threat)</td>
<td>$-88.42$</td>
<td>6</td>
<td>$\chi^2(6, N = 208) = 3.32$, $p &lt; .01$</td>
<td>$\chi^2(1, N = 208) = 3.32$, $p &lt; .01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>(type of task [PS/CT], uncontrollability, social-evaluative threat)</td>
<td>$-88.42$</td>
<td>9</td>
<td>$\chi^2(9, N = 208) = 3.32$, $p &lt; .01$</td>
<td>$\chi^2(1, N = 208) = 3.32$, $p &lt; .01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>(type of task [PS/CT], uncontrollability, timing of assessment, time of day)</td>
<td>$-88.42$</td>
<td>10</td>
<td>$\chi^2(10, N = 208) = 3.32$, $p &lt; .01$</td>
<td>$\chi^2(1, N = 208) = 3.32$, $p &lt; .01$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The log-likelihood ratio indicates the degree to which the predicted model deviates from the data. A significant chi-square indicates the addition of predictors improves the model fit. Bold typeface signifies the model with the best, most parsimonious fit to the data. Dashes indicate that the models were not compared. PS/CT = public speaking/cognitive combination task versus other tasks contrast.

Commentary:

- **Uncontrollability.** Consistent with the hypothesis that uncontrollable contexts would elicit greater cortisol changes than those that were controllable, uncontrollability significantly predicted the effect sizes ($\gamma_{01} = 0.28$, $SEM = 0.065$), $t(205) = 4.35$, $p < .01$. Both controllable and uncontrollable tasks elicited a significant cortisol response; however, the uncontrollable tasks were associated with a significantly greater cortisol response ($d = 0.52$, $CI = 0.38, 0.66$, $p < .01$) than the controllable tasks ($d = 0.16$, $CI = 0.03, 0.29$, $p < .05$).

- **Social-evaluative threat and uncontrollability.** Because both social-evaluative threat and uncontrollability significantly predicted the effect sizes, we then examined the relationship between these constructs. First, to examine whether social-evaluative threat and uncontrollability were independent predictors of the effect sizes, we simultaneously entered social-evaluative threat and uncontrollability (with the methodological controls) into a regression equation. Both social-evaluative threat ($\gamma_{01} = 0.34$, $SEM = 0.073$), $t(204) = 4.64$, $p < .01$, and uncontrollability ($\gamma_{02} = 0.36$, $SEM = 0.063$), $t(204) = 5.67$, $p < .01$, significantly predicted the effect sizes, with methodological factors controlled for. There was not a difference in predictive strength between social-evaluative threat and uncontrollability ($p > .20$). This demonstrates that both social-evaluative threat and uncontrollability are significant predictors of the cortisol response, independent of each other. Social-

---

11 Only two studies had three elements of uncontrollability. The small number of studies with a rating of 3 precluded analyzing them as a discrete category, and they were included in a group along with those receiving a rating of 2.
evaluative threat and uncontrollability together explained 26% of the variance after the time of day was accounted for. This was a significant increase in explained variance beyond time of day alone, $F(2, 204) = 34.65, p < .01$. Together, time of day, social-evaluative threat, and uncontrollability accounted for 32% of the between-studies variance in effect sizes.

We next addressed the additive effects of social-evaluative threat and uncontrollability and examined the role of task performance in eliciting cortisol responses. We proposed that motivated performance situations, or active performance tasks with the potential for evaluation along a self-relevant domain, coupled with uncontrollability and/or social-evaluative threat, would provide a context capable of triggering large changes in this system. To test this, we separated the studies into five categories: (a) motivated performance tasks with both social-evaluative threat and uncontrollability (e.g., public speaking and mental arithmetic with time constraints, all performed before an evaluative audience), (b) motivated performance tasks with only uncontrollability (e.g., mental arithmetic with uncontrollable noise or time constraints), (c) motivated performance tasks with only social-evaluative threat (e.g., delivering a speech in front of an audience), (d) motivated performance tasks alone (without uncontrollability and social-evaluative threat, e.g., solvable anagrams without time limits), and (e) passive tasks (e.g., watching a film).

As shown in Figure 3, there was substantial variability in the effect sizes associated with these stressor categories. The passive tasks ($-0.07$) and motivated performance tasks (without social-evaluative threat or uncontrollability; $-0.08$) did not elicit significant cortisol responses ($CI = -0.36, 0.22, p > .20$; $CI = -0.27, 0.12, p > .20$, respectively). These were significantly different than the effect sizes obtained with tasks with at least one of these elements ($\gamma_{01} = 0.20, SEM = 0.049$), $t(205) = 3.98, p < .01$. Motivated performance tasks with uncontrollability and/or social-evaluative threat elicited significant cortisol responses; effect sizes for these tasks were all significantly different from zero ($p < .01$).

The motivated performance tasks with both social evaluation and uncontrollability had the largest effect size of 0.92 ($CI = 0.70, 1.14$), which was significantly greater than the effect sizes for motivated performance with only uncontrollability (0.32, $CI = 0.09, 0.55$) or only social-evaluative threat (0.35, $CI = 0.15, 0.54$; $\gamma_{03} = 0.16, SEM = 0.040$), $t(205) = 4.13, p < .01$. The effect sizes for motivated performance with only uncontrollability or only social evaluation were not different from each other ($p > .20$). Together, these analyses indicate that only motivated performance tasks with uncontrollability and/or social-evaluative threat elicit a significant cortisol response, indicating that task performance alone is not sufficient to trigger this system. Furthermore, the combination of social evaluation in an uncontrollable performance context is associated with the greatest cortisol changes. In fact, the effect size for this combination was quite high and almost 3 times that of the components separately.

Stressor task category, social-evaluative threat, and uncontrollability. It is possible that the type of stressor task confounds the strong relationship between social-evaluative threat, uncontrollability, and cortisol responses, because the public speaking/cognitive combination task is associated with uncontrollability ($r = .29, p < .01$) and high levels of social-evaluative threat ($r = .54, p < .01$). To rule out this alternative explanation, we simultaneously entered social-evaluative threat, uncontrollability, the public speaking/cognitive task contrast (public speaking/cognitive task vs. all other tasks), and the methodology controls into a regression equation. Uncontrollability and social-evaluative threat remained significant when type of task was controlled for: social-evaluative threat ($\gamma_{01} = 0.19, SEM = 0.042$), $t(203) = 4.52, p < .01$; uncontrollability ($\gamma_{02} = 0.31, SEM = 0.072$), $t(203) = 4.25, p < .01$. However, type of task was no longer significant ($\gamma_{03} = 0.011, SEM = 0.0080$), $t(203) = 1.31, p > .19$). In addition, social-evaluative threat and uncontrollability were significantly stronger predictors than type of task, $\chi^2(1, N = 208) = 13.61, p < .01$. The relationship between social evaluation, uncontrollability, and cor-

**Figure 2.** Mean ($\pm$ SEM) cortisol effect size ($d$) for studies using cognitive tasks, public speaking/verbal interaction tasks, public speaking/cognitive combination tasks, noise exposure, and emotion induction. ***$p < .001$.**
The addition of social-evaluative threat and uncontrollability model with the methodological controls alone (see Table 4, Row 3). The addition of cortisol assessment. Sizes, above and beyond that explained by time of day and timing of cortisol assessment. This provides evidence that social-evaluative threat and uncontrollability mediate the relationship between stressor and cortisol responses. The difference in effect size obtained with the public speaking/cognitive task combination compared with the other task categories can be explained through social-evaluative threat and uncontrollability.

**Model Comparison**

We next compared the model including social-evaluative threat, uncontrollability, and the methodological predictors with the model with the methodological controls alone (see Table 4, Row 4). The addition of social-evaluative threat and uncontrollability significantly improved the overall fit of the model (when tested individually or together). This is consistent with the previous analyses that found social-evaluative threat and uncontrollability significantly predicted effect sizes when the methodological predictors were controlled for.

The regression analyses indicated that social-evaluative threat and uncontrollability were stronger predictors of effect size than the type of task. To confirm this, we compared a full model (including social-evaluative threat, uncontrollability, type of task, and methodological controls) with several reduced models to determine the one that provided the best, most parsimonious fit to the data. The full model provided a better fit than one that included only the type of task and methodological controls (see Table 4, Row 5, Column 6). This reduction in fit when social-evaluative threat and uncontrollability are omitted from the model indicates the importance of these variables for explaining the pattern of data. However, there were no differences in fit between the full model and one that included only social-evaluative threat, uncontrollability, and the methodological controls (omitting type of task; see Table 4, Row 5, Column 7). Together, these analyses demonstrate that social-evaluative threat and uncontrollability are necessary to explain the pattern of effect sizes across studies; however, type of task does not provide additional information beyond that of social-evaluative threat and uncontrollability. Including the time of day, timing of cortisol assessment, social-evaluative threat, and uncontrollability provides the best, most parsimonious model for predicting effect sizes.

**Predicting Cortisol Recovery**

The preceding analyses strongly demonstrate that performance tasks characterized by uncontrollability and/or social-evaluative threat are associated with larger average effect sizes during the 1-hr period after the stressor onset than other types of tasks. However, this reflects differences in the overall magnitude of the cortisol changes and does not specifically address whether these tasks differentially affect recovery, or the degree to which elevations persist after the stressor has ended (Linden et al., 1997). Because some evidence suggests that different factors may influence these processes (e.g., Matthews et al., 2001), we tested whether uncontrollability and social-evaluative threat also pre-
dicted patterns of recovery. Using the effect sizes generated from samples obtained 0–60 min from stressor termination, we first established the time course for poststressor cortisol declines across all of the studies. We then tested whether uncontrollable performance tasks with social evaluative threat showed delays in recovery compared with tasks without these characteristics (controlling for time of day).

Across all of the studies, cortisol levels decreased as time elapsed from the end of the stressor; the linear effect for time was significant (\( \gamma_{10} = -0.0083, SEM = 0.0011 \)), \( t(196) = -7.52, p < .01 \). Figure 4 displays the effect sizes for the samples obtained 0–20 min, 21–40 min, and 41–60 min after stressor termination. Cortisol levels were nearly twice as high 0–20 min poststressor (\( d = 0.38, CI = 0.28, 0.48, p < .01 \)) as in the subsequent 21–40-min period (\( d = 0.26, CI = 0.13, 0.39, p < .01 \)), and levels continued to decline 41–60 min poststressor (\( d = -0.05, CI = -0.19, 0.09, p > .20 \)). Orthogonal contrasts revealed that the differences between time intervals were significant: 0–20 min versus 21–60 min (\( \gamma_{01} = 0.14, SEM = 0.0177 \)), \( t(196) = 8.35, p < .01 \); 21–40 min versus 41–60 min (\( \gamma_{02} = 0.20, SEM = 0.0330 \)), \( t(196) = 6.12, p < .01 \). Whereas the 0–20-min and 21–40–min periods were associated with significant cortisol responses, the effect size obtained 41–60 min from stressor termination was not significant. This indicates that, overall, cortisol levels return to prestressor levels by 41–60 min after the end of the stressor.

To examine the relationship between social-evaluative threat, uncontrollability, and cortisol changes after stressor termination, we divided the tasks into (a) motivated performance tasks with both social-evaluative threat and uncontrollability, (b) motivated performance tasks with either social-evaluative threat or uncontrollability, and (c) motivated performance tasks without social-evaluative threat or uncontrollability and passive tasks (because there were significant differences between these groups for the average effect size analyses; see Figure 3). The effect sizes for these three categories during each 20-min poststressor interval are displayed in Figure 5.

During the 0–20-min poststressor period (Figure 5, Column 1), performance tasks with social-evaluative threat and uncontrollability (\( d = 0.85, CI = 0.63, 1.07, p < .01 \)) and performance tasks with one of these components (\( d = 0.25, CI = 0.14, 0.36, p < .01 \)) were associated with significant cortisol responses; the effect sizes for these categories were significantly greater than zero. However, the group of passive tasks and performance tasks without either component were not associated with significant cortisol changes (\( d = -0.02, CI = -0.15, 0.11, p > .20 \)). The effect size of 0.85 for the performance tasks with social-evaluative threat and uncontrollability was significantly different from the other task categories (\( \gamma_{01} = 0.27, SEM = 0.0361 \)), \( t(186) = 7.53, p < .01 \). There was also a difference in effect sizes for the tasks with either social-evaluative threat or uncontrollability and passive tasks/performance tasks without either component (\( \gamma_{01} = 0.13, SEM = 0.0462 \)), \( t(186) = 2.91, p < .01 \). These analyses demonstrate that tasks with social-evaluative threat and/or uncontrollability were associated with significant cortisol elevations 0–20 min poststressor, and the tasks with both components had an effect size nearly 3 times the size of tasks with either component alone.

During the 21–40-min poststressor period (Figure 5, Column 2), the uncontrollable performance tasks with social-evaluative threat still exhibited a strong, reliable cortisol effect (\( d = 0.74, CI = 0.49, 0.99, p < .01 \)). However, the performance tasks with either uncontrollability or social-evaluative threat were no longer associated with significant cortisol changes during this poststressor time interval (\( d = 0.08, CI = -0.07, 0.23, p > .20 \)), and the effect size associated with passive tasks and/or tasks without either component again was not significantly different from zero (\( d = -0.14, CI = -0.45, -0.17, p > .20 \)). This demonstrates that whereas the tasks with both social-evaluative threat and uncontrollability still showed significant elevations 21–40 min poststressor, the cortisol levels associated with all of the other tasks were at baseline levels. The difference in effect size between uncontrollable, social-evaluative tasks and the other categories is highly significant (\( \gamma_{01} = 0.26, SEM = 0.0450 \)), \( t(89) = 5.81, p < .01 \); the effect size for tasks with both components was almost 7 times larger than that for any other group. There were no significant differences between the combined category of single-component performance tasks and passive tasks and performance tasks without either component (\( \gamma_{01} = 0.13, SEM = 0.0950 \)), \( t(89) = 1.36, p > .17 \).

Despite small numbers of studies that assessed cortisol 41–60 min poststressor (Figure 5, Column 3), a similar pattern of results...
emerged during this interval. Again, performance tasks with both social-evaluative threat and uncontrollability were the only type to show persistent cortisol elevations ($d = 0.28$, CI = 0.03, 0.53, $p < .05$). The other task groups were not associated with significant cortisol changes ($d = -0.21$, CI = $-0.92, 0.50$, $p > .20$; and $d = -0.08$, CI = $-0.31, 0.16$, $p > .20$). The effect size for uncontrollable, social-evaluative tasks was significantly different than those for the other types of tasks (social-evaluative threat and uncontrollability was significantly different than all other tasks) as predictors of effect sizes 21–40 min and 41–60 min from the end of stressor. The social-evaluative, uncontrollable contrast for recovery times was no longer significant ($ps > .20$), whereas peak response remained a strong predictor of the effect sizes ($p < .01$). Therefore, this provides evidence that the persistent cortisol elevations observed after social-evaluative, uncontrollable tasks are primarily due to their larger peak response.

**Predicting ACTH Responses**

Although social-evaluative, uncontrollable conditions are associated with greater cortisol responses, it is unclear whether ACTH would respond in the same manner. We examined this issue on a subsample of 39 studies that also assessed ACTH responses to acute laboratory stressors. First, we conducted analyses to test whether ACTH responses would correlate with cortisol responses. Then, we tested whether uncontrollable, social-evaluative tasks

---

13 When several cortisol samples were taken during this time period, the effect sizes were averaged to obtain a single peak value.

14 Because of the substantial overlap between effect sizes obtained 21–40 min from stressor onset and 0–20 min from stressor termination, we did not examine this first recovery period in these analyses.
elicited greater ACTH changes compared to those without these characteristics.

It is important to note that these analyses were conducted on a nonrandom sample of studies in the meta-analysis. For example, they are associated with a larger overall cortisol response; the average effect size for cortisol in this subsample is 0.59 (CI = 0.37, 0.81, p < .01), which is considerably higher than the overall effect of 0.31 obtained in the sample as a whole (probably because these studies are more likely to have social-evaluative threat), $\chi^2(1, N = 208) = 5.18, p < .05$. Therefore, the overall effect size for ACTH responses obtained with this subsample of studies may not be representative of the meta-analytic sample as a whole (had these responses been assessed). However, our primary interest is one of comparison: comparing patterns of ACTH and cortisol responses, and comparing ACTH responses to uncontrollable, social-evaluative tasks with responses to other tasks without these characteristics.

ACTH responses were a significant predictor of cortisol responses ($\gamma_{10} = 0.60$, $SEM = 0.13$), $t(38) = 4.72, p < .01$. This strong association between ACTH and cortisol is illustrated in Figure 6, which shows the timecourse of ACTH and cortisol responses in this sample in 10-min epochs from stressor onset. The ACTH response peaks approximately 11–20 min from stressor onset ($\gamma_{10} = -0.67, SEM = 0.27$, $t(38) = -2.50, p < .05$), and shows a linear decline over time ($\gamma_{10} = -0.025, SEM = 0.069$, $t(38) = -3.66, p < .01$). ACTH responses precede cortisol by approximately 10–20 min.

We then tested whether uncontrollable, social-evaluative conditions were associated with larger ACTH responses compared to tasks without both of these characteristics (motivated performance tasks with either social-evaluative threat or uncontrollability, motivated performance tasks without either component, and passive tasks). As shown in Figure 7, uncontrollable social-evaluative stressors were associated with greater ACTH responses compared with the other tasks ($\gamma_{10} = 0.48, SEM = 0.19$, $t(37) = 2.53, p < .05$). The ACTH and cortisol responses are strikingly parallel in this sample of studies. Together, these analyses demonstrate that among healthy adults, there appears to be a tight coupling between these two components of the HPA axis; both ACTH and cortisol show greater responsivity to uncontrollable, social-evaluative stressors.

**Supplemental Analyses**

**Distress and Cortisol Responses**

It is possible that social-evaluative threat and uncontrollability are more likely to elicit cortisol responses because these conditions are simply more distressing than the other tasks. In other words, it is not specific features of the tasks that lead to heightened cortisol responses (i.e., social-evaluation), but rather that these tasks are more stressful in general and therefore more likely to elicit feelings of distress, which are in turn related to increases in cortisol. To pursue this question, we identified the subsample of 69 studies that assessed self-reported psychological states (e.g., perceived stress, negative affect, arousal) both pre- and poststressor. We then computed effect sizes for the negative subjective responses, using the same procedures as the cortisol calculations. There was variability in the negative subjective states assessed; most measured stress or general negative affect (e.g., stress, distress, general negative affect, anxiety; $\kappa = 52$), but others assessed more somatic responses, such as tension or arousal ($\kappa = 17$).

Overall, the stressors elicited negative subjective responses ($d = 1.04, CI = 0.76, 1.27, p < .01$), indicating that the tasks were experienced as distressing. We then tested whether the social-evaluative and/or uncontrollable tasks led to greater increases in negative psychological states compared to those without these characteristics. There were no differences in subjective distress between tasks with and without social-evaluative threat, $t(67) = -0.05, p > .20$, or with and without uncontrollability, $t(67) = 1.33, p > .18$. In addition, there were no differences in subjective distress between tasks with and without uncontrollability, $t(67) = -0.05, p > .20$, or with and without uncontrollability, $t(67) = 1.33, p > .18$. In addition, there were no differences in subjective distress between tasks with and without regret, $t(67) = -0.20, p > .80$, or with and without regret, $t(67) = 0.27, p > .80$.

---

15 The small number of studies ($k = 3$) that used passive tasks or motivated performance tasks without either social-evaluative threat or uncontrollability precluded testing them as a separate category, and they were combined with the others.

16 When studies assessed psychological responses only after task completion or presented only differences between groups, effect sizes were unable to be calculated, so these studies were not included in the analyses. In studies in which multiple psychological responses were assessed, the most affective-based response was selected (e.g., selected negative affect over arousal). Finally, total negative emotion responses were used in effect size estimates if multiple subscales were reported (e.g., total negative emotion rather than anger, etc.).
distressing than the other task categories. This indicates that the stress or negative affective state assessments (excluding the results were obtained when the analyses were repeated with only the stress or are perceived as more stressful. 

Social-evaluative threat and uncontrollability are not associated with cortisol solely because they induce more psychological distress between the uncontrollable, social-evaluative tasks and those in the other categories, $t(67) = 0.47, p > .20$. Identical results were obtained when the analyses were repeated with only the stress or negative affective state assessments (excluding the more somatic tension and arousal measures). This indicates that social-evaluative, uncontrollable stressors are not simply more distressing than the other task categories. Furthermore, there was not a relationship between cortisol and distress ($\gamma_{01} = 0.064, SEM = 0.067), t(66) = 0.94, p > .20$; there was no evidence that stressors that induced greater levels of general negative affect or distress were associated with greater cortisol responses. Again, excluding the studies that assessed more somatic responses did not change the results ($\gamma_{01} = 0.074, SEM = .071), t(48) = 1.01, p > .20$. Together, these analyses suggest that social-evaluative threat and uncontrollability are not associated with cortisol solely because they induce more psychological distress or are perceived as more stressful.

**Potential Research Group Effects**

The laboratory of Clemens Kirschbaum and Dirk Hellhammer contributed 19 of the 208 studies (9%) included in the meta-analysis, and many of these studies elicited significant cortisol responses. To exclude the possibility that experimenter effects were driving the results of the meta-analysis, we tested whether their studies were associated with greater effect sizes compared with studies from other laboratories. Kirschbaum–Hellhammer authorship was a significant predictor of effect sizes ($\gamma_{01} = 0.25, 0.052), t(206) = 4.93, p < .01$. However, the effect sizes may be higher in their studies compared with others because their studies were more likely to be conducted in the afternoon. $\chi^2(2, N = 208) = 6.92, p < .05$, and to have uncontrollability, $\chi^2(1, N = 208) = 11.79, p < .01$, and high levels of social-evaluative threat, $\chi^2(2, N = 208) = 52.8, p < .01$, and each of these characteristics was associated with greater cortisol responses. When these factors were controlled for, Kirschbaum–Hellhammer authorship was no longer a significant predictor of the effect sizes ($p > .20$). This indicates that the strong cortisol responses in their lab are more a product of the procedures and type of stressor tasks that they used rather than an idiosyncratic laboratory effect.

To further test the contribution of the Kirschbaum–Hellhammer studies to the meta-analysis, we reconducted the primary theoretical analyses, excluding the 19 studies conducted in their laboratory. Without these studies, social-evaluative threat and uncontrollability remained highly significant predictors of cortisol changes ($ps < .01$). The biggest concern related to the effects involving the uncontrollable, social-evaluative tasks, because the Kirschbaum–Hellhammer laboratory contributed 18 (out of 30) studies to this category. However, the average effect size of the uncontrollable, social-evaluative tasks, excluding their studies, was 0.88; this is quite similar to the average effect size when their studies are included (0.92). Furthermore, contrast analyses comparing the tasks with both uncontrollability and social-evaluative threat to those with only one of these components remained significant ($p < .05$). Together, these analyses provide quite strong evidence that the studies conducted in the Kirschbaum–Hellhammer laboratory are not driving the meta-analytic results.

**Publication Bias**

We conducted additional analyses to determine whether there could be a potential publication bias in this literature; this can be problematic in meta-analysis because the likelihood of publication could increase with positive findings and/or stronger effects (although this plagues traditional narrative reviews as well; Rosenthal, 1991). As the meta-analysis included both unpublished dissertations as well as published empirical reports, we tested whether publication status was a significant predictor of effect sizes. Publication status was not a significant predictor of effect sizes ($\gamma_{01} = -0.12, SEM = 0.17), t(205) = -0.74, p > .20$, indicating that patterns of cortisol changes in unpublished dissertations were similar to those found in the published literature.

Furthermore, we examined the relationship between the effect sizes and design features of the studies that could increase the likelihood of publication (Begg, 1994). Among the published studies, the sample size and the presence of methodological features meant to reduce error variance (i.e., constant time of day, health screens) were not significant predictors of the effect sizes ($ps > .20$). This suggests there is not a significant publication bias in this literature. This could be because many of the studies assessed other biological parameters (e.g., immunologic or sympathetic products or other neuroendocrine hormones) in addition to cortisol. Studies that did not find a cortisol effect in response to acute stressors often reported changes in these other outcomes, which could lead to successful publication despite the nonsignificant cortisol results.

**Discussion**

This research synthesis reviews 208 studies of plasma or salivary cortisol responses to acute psychological laboratory stressors in healthy adults. Using a multilevel modeling analytic approach, we found that such stressors significantly increase cortisol levels. The findings provide evidence that, overall, like physical stressors (e.g., electric shock), psychological stressors can activate the HPA.
axis, on average increasing cortisol levels 0.31 standard deviations above baseline values in these studies. This overall effect size of 0.31 is considered small (J. Cohen, 1988). However, there was substantial variability in the effect sizes depending on the nature of the stressor task; some tasks provoked quite large cortisol responses, whereas others did not engage this system (effect sizes ranged from −0.08 to 0.92). Further analyses pointed to the stressor characteristics that account for this variability. As predicted, tasks that included social-evaluative threat, in which others could negatively judge performance, particularly when the outcome of the performance was uncontrollable, provoked larger and more reliable cortisol changes than stressors without these particular threats. These characteristics affect both the overall magnitude of the response and the recovery trajectory, as well as patterns of ACTH changes. Clearly, these findings refute the notion that all psychological stressors elicit cortisol responses. They also call into question the presence of a nonspecific physiological response to all stressors that include HPA activation (e.g., Selye, 1956; see Kemeny, 2003).

A great deal of debate exists about the relationship between stressful circumstances and HPA activation. Some have argued that most stressors elicit a cortisol response, whereas others propose that only stressors with certain characteristics have this capability (e.g., the outcome of the event is uncontrollable, the task is of long duration, the circumstances evoke distress). We found that cognitive tasks (e.g., mental arithmetic, Stroop), verbal interaction tasks (e.g., public speaking, interview), and public speaking/cognitive task combinations elicited significant cortisol responses. However, noise exposure and emotion induction tasks (e.g., film) were not associated with significant elevations in cortisol levels.

These findings have several theoretical implications. First, they argue against the perspective that the experience of distress is a sufficient condition to elicit cortisol responses. Despite successfully producing negative affective states, the emotion induction studies, on average, failed to activate the cortisol system. Furthermore, we found no relationship between increases in subjective distress and/or negative affect in response to the stressors and cortisol changes. Second, the findings question the assumption that “stress” is a one-dimensional construct. Not all stressors are equivalent, leading to stereotyped physiological responses; only certain types of stressors, particularly those with specific characteristics, were associated with cortisol elevations. It is possible that the current tendency in the literature to use the word stress in a vague and diffuse way has prevented focused research on specific kinds of threats that can affect health-relevant physiological systems, impeding scientific discovery in this important area (Ader, 1980). These findings support the notion that “organisms meet these challenges and dangers by integrated behavioral, physiological patterns of response that are appropriate to the task” (Weiner, 1992, p. 33).

The public speaking/cognitive task combination was associated with greater cortisol responses than other types of stressors; in fact, the effect size associated with this type of task was nearly twice as large as the effect sizes for the other task categories. Prototypical public speaking/cognitive tasks included performing mental arithmetic with time constraints and delivering a speech in front of an evaluative audience. These tasks were not, on average, longer than the other types, and controlling for stressor length did not alter the results, suggesting that duration was not the important element driving the differential responses. Instead, we determined that this robust cortisol effect was due to the presence of two characteristics during motivated performance tasks: outcome uncontrollability and social-evaluative threat. In multiple analyses, uncontrollability and social-evaluative threat were stronger predictors of responsivity than the task category. Most important, the relationship between public speaking/cognitive task combinations and cortisol changes was mediated by the presence of social evaluation and outcome uncontrollability. Thus, consistent with our theoretical model, one important set of determinants of cortisol responses appears to consist of (a) a motivated performance task, (b) relative uncontrollability of task outcome, and (c) the presence of social evaluation.

Testing the Theoretical Model

Our overall premise is that threats to central goals, such as physical self-preservation, elicit cortisol responses as part of an adaptive mobilization of energy to reduce the threat. In addition to physical self-preservation, we propose that individuals are motivated to preserve the social self by maintaining social esteem, status, and acceptance. The social self could be threatened in situations where lack of a valued attribute or possession of an undesired quality could be revealed, leading to the loss of social esteem, respect, or social status. Drawing on the animal literature that demonstrates that threats to social status in dominance hierarchies elicit HPA activation (e.g., Sapolsky, 1993), we proposed that analogous social threats could trigger cortisol responses in humans as well. In the laboratory, situations characterized by social-evaluative threat, in which an important aspect of the self-identity is or could be negatively judged by others, could threaten this goal, particularly when there is the potential for poor performance in a domain that is valued by the group (e.g., competence, intelligence; Leary & Baumeister, 2000). Motivated performance situations, which are goal-relevant, active performance tasks with the potential for evaluation, provide such a context (e.g., mental arithmetic, public speaking; Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996). Motivated performance tasks coupled with social-evaluative threat could be potent elicitors of a cortisol response because the potential for exposed failure poses a significant threat to maintaining social esteem.

The threat could be further augmented in uncontrollable conditions where failure is imminent; uncontrollable situations create a context of forced failure in which it would be quite difficult to succeed despite one’s best efforts or impossible to avoid negative consequences (Henry & Grim, 1990; Weiner, 1992). The metaanalytic findings provide strong support for our theoretical model, which posits that the cortisol system is activated in goal-relevant situations (motivated performance tasks) when a central goal is saliently threatened (social-evaluative threat) and the process for attaining this goal is impeded (uncontrollability).

Social-Evaluative Threat

Tasks characterized by social-evaluative threat, in which an evaluative audience or negative social comparison was present or the performance was captured on a permanent record (e.g., videotape), were associated with an effect size of 0.67; this was over 3 times as large as the effect size for tasks without a social-
of shame and related emotions, and that the experience of this
appraisals generated under these conditions, lead to the experience
meny, in press).

Studies that included multiple social-evaluative components
(e.g., videotape and audience) showed greater cortisol elevations
than those that included just one element, demonstrating a graded
cortisol response to increases in social evaluation. However,
among the various study elements that were coded as inducing
social-evaluative threat, the presence of an audience or a negative
social comparison during the performance heightened cortisol
responses to a greater degree than did the presence of video record-
ing. This suggests that the real-time presence of evaluative oth-
ers—rather than the potential for others to evaluate the
performance in the future—is most important for eliciting cortisol
responses.

Why might the presence of evaluative others increase the effect
on cortisol? There are several potential explanations. First, indi-
viduals may behave differently during a task with social evalua-
tion. For example, they may exert more effort when performing the
task, and this effort may have physiological effects. However,
empirical studies have found that engagement and/or effort are
associated with activation of the sympathetic nervous system,
whereas cortisol is unrelated to this dimension (Buchanan, al’Absi,
& Lovallo, 1999; Lundberg & Frankenhuaser, 1980; Peters et al.,
1998). This suggests that increased effort does not explain the
relationship between social-evaluative threat and cortisol responses.

Second, individuals may appraise the situation and themselves
differently when being observed. Distinctive appraisals may not
increase psychological reactivity in a general way. Consistent with
this premise, we found that social-evaluative threat was not asso-
ciated with greater increases in self-reports of distress, and that
these more general emotional states were not correlated with
cortisol responses. Instead, a more specific set of cognitive and
emotional reactions may actually be produced under conditions of
social-evaluative threat.

Theoretical and empirical evidence suggests that self-related
cognitions and emotions are driven by the degree to which others
are rejecting of the self (Gilbert, 1997; Leary et al., 1995, 2001).
Thus, the presence of others in a context of potential failure could
induce self-evaluative states, which may have specific physiological
effects. Although there is little prior evidence linking self-
evaluative states and the cortisol system, both longitudinal and
experimental studies have shown that negative self-related cogni-
tive appraisals have specific immunological correlates, supporting
the existence of distinctive neurophysiological associations (Cole,
Kemeny, & Taylor, 1997; Dickerson, Kemeny, Aziz, Kim, & Fahey,
2004; Kemeny & Dean, 1995; T. T. Lewis, Kemeny, Myers, & Wyatt,
2004; Segerstrom, Taylor, Kemeny, Reed, & Visscher, 1996; for review, see Dickerson, Gruenewald, & Ke-
meny, in press).

We have proposed that social-evaluative threat, and the self-
appraisals generated under these conditions, lead to the experience
of shame and related emotions, and that the experience of this
family of emotions could be tied to specific physiological changes
(Dickerson et al., in press). Consistent with this premise, self-

Uncontrollability

Because uncontrollability could impede progress toward attaining
goals, we predicted that uncontrollable stressors would elicit
greater cortisol responses than controllable ones. Despite repeated
claims that cortisol is elicited in uncontrollable conditions, few
empirical studies in humans have manipulated uncontrollability
(\(k = 6\)), and these have led to mixed effects (for review, see Peters
et al., 1998). However, across the 208 studies, uncontrollability
emerged as a significant predictor of the cortisol response; con-
trollable tasks were associated with an effect size of 0.16, whereas
uncontrollable tasks had an average effect size nearly 3 times
larger (0.52). Uncontrollable contexts contained elements that in-
formed participants they were failing or could not avoid negative
consequences, including manipulated task difficulty (impossible
tasks, time constraints), false feedback of poor performance, ha-
rassment, or the presence of auditory distraction or other emotion-
ally distressing stimuli when no behavioral methods for avoiding
the stimuli were possible. Animal studies have documented dif-
ferential HPA activation in uncontrollable versus controllable con-
ditions (Davis et al., 1977; Dess et al., 1983; Hanson et al., 1976;
Swenson & Vogel, 1983), and these meta-analytic results extend
this literature to humans. Furthermore, the finding that uncontrol-
lability leads to increased cortisol activation adds to the evidence
that loss of control can have negative effects on psychological,
physiological, and health outcomes (e.g., Chorpita & Barlow,
1998; Peterson, Maier, & Seligman, 1993).

However, uncontrollability was not unilaterally associated with
cortisol responses. The outcome uncontrollability had to occur in
the context of a motivated performance task in these studies to
elicit cortisol changes. For example, uncontrollable passive situa-
tions (e.g., noise exposure) did not significantly increase cortisol
levels. This suggests that uncontrollability must threaten an im-
portant motivational domain to trigger this system; past failures to
document uncontrollability—cortisol associations in human studies
may have resulted from the fact that many of the tasks selected did
not pose a substantial threat to a central goal. This reasoning is
supported by animal studies in which physical integrity is threat-
ened while uncontrollability is manipulated (e.g., receiving con-
trollable vs. uncontrollable electric shock), resulting in differential
HPA activation. In other words, these data suggest that the expe-
rience of not having control over a situation does not necessarily
elicit this physiological response. However, being in a situation in
which an important goal is threatened, when the desired outcome
is not contingent on the organism’s behavior, appears to trigger
cortisol activation.
Uncontrollable Social-Evaluative Threat

We predicted that social-evaluative contexts with uncontrollable outcomes would elicit the largest cortisol increases. Consistent with this premise, the meta-analysis demonstrated that uncontrollable motivated performance tasks performed in the presence of others are the strongest elicitors of cortisol activation ($d = 0.92$; considered a very large effect; J. Cohen, 1988). Together, uncontrollability and social-evaluative threat accounted for 26% of the between-studies variance in effect sizes. Similar findings of uncontrollability augmenting cortisol responses to social status-related threats have been reported in the animal literature. Low status on the social hierarchy is typically associated with elevated cortisol levels, and activation of this system is further heightened among subordinate baboons facing uncontrollable stressors (Sapolsky, 1993).

However, motivated performance situations without either of these components (uncontrollability, social-evaluative threat) did not elicit significant cortisol responses, indicating that performing a difficult task by itself is not enough to trigger this system. Motivated performance tasks without social-evaluative threat and/or uncontrollability might not pose a substantial threat to maintaining the social self, because there is no potential for poor performance to be exposed to others (in contrast to social-evaluative situations), and success is more likely (in contrast to uncontrollable situations). These conditions could result in states of "challenge," occurring when resources exceed situational demands (Lazarus & Folkman, 1984), which have not been theorized to result in HPA activation (Blascovich & Tomaka, 1996; Dienstbier, 1989).

We proposed that motivated performance situations without social-evaluative threat would not pose a direct threat to the goal of maintaining the social self; social others could not explicitly evaluate the participant, and, therefore, such situations would not activate the HPA. Results of the meta-analysis indicated that motivated performance tasks without social-evaluative threat or uncontrollability did not elicit a significant cortisol effect ($d = -0.06$). However, motivated performance tasks without social-evaluative threat that were uncontrollable did elicit small, but significant, cortisol changes ($d = 0.32$). In motivated performance tasks without social-evaluative threat, there is no evaluative audience and no permanent record of performance. However, the experimenter is present, so there is still the potential that the task performance could be evaluated. Although this more oblique threat does not appear to be sufficient to activate this system under uncontrollable conditions, the context of forced failure in uncontrollable motivated performance tasks could make even this minimal form of assessment salient, leading to cortisol elevations. It is unclear whether uncontrollable performance tasks under completely private conditions, in which only the participant would know the outcome of the performance, would elicit cortisol responses. We have shown that revealing negative aspects of the self in an anonymous and confidential setting does not elevate cortisol levels (Dickerson et al., 2004), suggesting that at least the potential for critique may be necessary to elicit a cortisol response.

Predictors of Recovery

We found that uncontrollable, social-evaluative conditions not only affected overall cortisol responses, but also influenced the recovery process, or the degree to which cortisol elevations persist after stressor termination (Linden et al., 1997). Only exposure to the uncontrollable, social-evaluative performance tasks resulted in a failure to return to baseline cortisol levels within 1 hr after the end of the stressor. During the 41–60 min poststressor interval, these tasks were still associated with substantial cortisol elevations of 0.28 standard deviations above prestressor levels. Among the tasks with only one of these components, cortisol levels had returned to baseline by 21–40 min poststressor (despite a significant cortisol response during the 0–20 min poststressor period). Therefore, the cortisol changes associated with uncontrollable, social-evaluative tasks persisted at least 40 min longer than those associated with other types of tasks. Most research has focused on the overall magnitude of cortisol change in response to psychological stressors; very few studies have examined how different contexts or individual difference factors predict cortisol recovery (cf. Earle et al., 1999; Matthews et al., 2001; Roy et al., 2001). However, the meta-analytic findings suggest that elucidating the factors associated with recovery processes could be a fruitful avenue of inquiry.

In this acute, time-limited context, the delayed recovery associated with the uncontrollable, social-evaluative performance tasks was closely tied to greater overall peak cortisol changes; the peak response mediated the relationship between these tasks and longer times to return to baseline. In other words, the differences in recovery appeared to be driven by the greater peak levels elicited by the uncontrollable, social-evaluative tasks. Thus, it is possible that delayed recovery in this context may be due to slower clearance of the high levels of cortisol from circulation. On the other hand, whereas a significant portion of the variance in recovery may be explained by the magnitude of the peak response, part may also be due to the extent to which activation persists after stressor termination. Research is needed to determine the extent to which the prediction of recovery is dependent on the peak response.

Several critical questions remain regarding the relationship between stressors and cortisol recovery. First, it will be important to examine how recovery processes may play out in naturally occurring stressors. Acute laboratory contexts are time-limited, and participants are debriefed; rarely would real-life stressors have such a circumscribed end. Cognitive and affective processes that lead to HPA activation may be more likely to be extended in naturalistic stressors, which could lead to greater persistence in cortisol elevations. Even in the laboratory context, there were not enough studies with recovery periods beyond 60 min to determine the point at which effects would return to baseline. Second, it is important to determine whether peak cortisol change and recovery are related or relatively independent processes in a naturalistic, non-time-limited context. For example, some studies have found a rapid, strong activation of the HPA axis coupled with a rapid recovery in a subset of individuals. In what contexts and in whom is this pattern likely to occur? Finally, future research should focus on elucidating the unique mechanisms and predictors of recovery processes.

ACTH Responses

ACTH is released from the anterior pituitary and is the precursor to cortisol, as it stimulates the release of cortisol from the adrenal cortex. Some studies have found increases in both ACTH and
cortisol in response to acute psychological stressors (e.g., Kirschbaum et al., 1999), whereas others have found increases in ACTH, but not in cortisol (e.g., Cacioppo et al., 1995; Malarkey, Kiecolt-Glaser, Pearle, & Glaser, 1994; van der Pompe et al., 1996). Therefore, it is unclear to what degree certain stressors may elicit joint increases in ACTH and cortisol, particularly given evidence that certain psychological states, such as depression, can lead to dysregulations at different levels of the HPA axis (for review, see Gold, Licinio, Wong, & Chrousos, 1995). In a subsample of 39 studies that assessed ACTH responses, the effect sizes for ACTH were highly correlated with the effect sizes for cortisol, suggesting that both were often stimulated at a similar level. Very little research has examined how levels of ACTH correlate with levels of cortisol produced. Second, the time course of the hormones were not equivalent, with ACTH peaking approximately 10–20 min before cortisol. This issue of timing could be one reason a dissociation has been commonly reported in the literature; as these hormones peak at different times, repeated assessments may be necessary to capture maximal changes in both products.

We demonstrated that social-evaluative, uncontrollable stressors were associated with greater increases in ACTH responses compared to stressors without these elements. Again, we observed a close association between ACTH and cortisol response to these stressors, with the overall effect sizes for both hormones nearly equivalent. Together, these findings suggest that, among healthy individuals, there does appear to be a strong association between ACTH and cortisol. Furthermore, like cortisol, ACTH does seem to be preferentially activated under uncontrollable, social-evaluative conditions.

Health Implications

Theoretical models have proposed that chronically experiencing conditions that elicit HPA activation could lead to a wide array of negative physiological changes that can have long-term health effects (e.g., Dienstbier, 1989; McEwen, 1998). This could occur through several pathways, including repeated activation of the HPA system as a result of frequent exposure to stressful conditions or a failure to shut down the response after stressor termination (i.e., lack of recovery). The consequence of this prolonged exposure to stress hormones has been called allostatic load, or a cumulative toll on the body resulting from chronic overactivation of the stress system (McEwen, 1998; McEwen & Stellar, 1993). These response patterns are thought to increase the risk of a number of negative health outcomes, such as diabetes, hypertension, cancer, and cardiovascular disease (McEwen, 1998).

A variety of stressful circumstances have been thought to increase allostatic load; we build on this model by defining specific types of conditions that activate the HPA system. Because uncontrollable conditions that threaten the social self are potent elicitors of cortisol, these situations (if experienced repeatedly) could be particularly likely to contribute to allostatic load. Studies in primates have found that chronic, uncontrollable social threats resulting from subordinate status can lead to a number of negative biological outcomes associated with allostatic load (e.g., decreased lymphocyte counts, high cholesterol levels, development of atherosclerosis; Kaplan, Manuck, Clarkson, Lusso, & Taub, 1982; Sapolsky, 1993), and significant relationships have emerged between cortisol levels and some of these parameters (Sapolsky, 1993). Research needs to document the effects of chronic forms of uncontrollable social threat and their impact on the HPA system and health in humans. However, we would predict that individuals who possess uncontrollable characteristics that result in social rejection (e.g., those who are stigmatized) could persistently experience uncontrollable social-evaluative threat; this could lead to chronic activation of this physiological system with its possible health implications. In addition, individuals who persistently appraise their social world as rejecting or who are particularly sensitive to rejection could also demonstrate these health vulnerabilities. Consistent with this premise, our research has demonstrated that rejection sensitivity in a stigmatized group of HIV-positive gay men predicts long-term immunologic and virologic alterations and increases in mortality (Cole et al., 1997; Cole, Kemeny, Fahey, Zack, & Naliboff, 2003).

The research synthesis demonstrated that social-evaluative, uncontrollable conditions resulted in greater cortisol changes as well as a delayed time to recovery; both have been discussed in the allostatic load framework. Some have theorized that examining patterns of cortisol recovery after stressors may actually be more health relevant (e.g., Dienstbier, 1989; Linden et al., 1997), and that pathology ensues when recovery is prevented from occurring (Sapolsky et al., 2000). Indeed, a quick, strong HPA response coupled with rapid recovery in many cases would be adaptive, providing the organism with the necessary energy to reduce the goal threat (Linden et al., 1997; Sapolsky et al., 2000). However, a sluggish return to baseline could result in longer overall exposure to stress hormones as well as indicate underlying dysregulation in the stress-responsive systems (Sapolsky et al., 2000). Thus, repeated or chronic exposure to social-evaluative, uncontrollable contexts may be particularly pathogenic because of the protracted recovery phase that may be associated with these conditions.

Methodological Implications

Methodological or procedural factors could preclude eliciting and/or capturing the cortisol response to acute stressors, which might have contributed to the variability in this literature. Although it has long been known that certain methodological factors can influence cortisol changes after psychological stressors (e.g., Mason, 1968; Kirschbaum & Hellhammer, 1994; Lovallo & Thomas, 2000), the conclusions and recommendations for the best laboratory procedures to utilize have been based on data from select sets of studies, which may not be representative of the literature as a whole. The research synthesis allows us to test the effects of these methodological factors on cortisol changes across a wide range of tasks and laboratories. By controlling for the methodological factors related to cortisol responses, we could clearly test the primary theoretical hypotheses without these confounds.

These meta-analytic results provide a blueprint for designing laboratory procedures to capture maximal cortisol changes in response to psychological stressors. First, cortisol assessments 21–40 min from stressor onset should be obtained. These samples are associated with the largest effect sizes (d = 0.38–0.41), whereas those from samples obtained at other time intervals (both less than 20 min and more than 40 min) are significantly smaller (d = 0.13–0.29). Given that the peak cortisol response consistently occurs during this 21–40-min window across the studies,
assessments obtained at other time points may miss maximal cortisol changes. In addition, it is important to measure the trajectory for recovery. These results suggest that samples obtained 60 min or longer after stressor termination may be optimal to chart the recovery process.

Second, studies of acute stressor effects should be conducted in the afternoon or utilize the proper time-of-day controls. Studies conducted in the morning have an average effect size of only 0.14, whereas those in the afternoon have an average effect of 0.46. Cortisol has a circadian rhythm in which levels dramatically decrease in the morning and are relatively stable in the afternoon, making cortisol responses to acute stressors much easier to detect later in the day. Although many studies control for the diurnal variation by running all participants at the same time of day (e.g., all in the morning), these results suggest this might not be enough; studies conducted in the morning may miss important stress-related perturbations in the cortisol system. If studies are conducted in the morning, including a no-stressor control group or within-subjects designs to chart naturally decreasing cortisol levels could be critical for making appropriate conclusions about cortisol activity. However, these findings cannot be utilized for decisions regarding the best time of day for studies assessing basal cortisol levels.

In addition, the results demonstrate that a short-duration, public speaking/cognitive task combination is a clear, reliable way to elicit a substantial cortisol response (e.g., Trier Social Stress Task; Kirschbaum, Pirke, & Hellhammer, 1993). It is important to note that only a limited number of types of tasks have been utilized to date in these studies; other tasks may be as successful or more successful in eliciting cortisol activity. Because the stressor duration is not associated with cortisol responses, this argues against the utility of extended stressor challenges. Shorter stressors, with the proper eliciting conditions, are equally as effective as longer tasks in increasing cortisol levels. However, simply utilizing a public speaking/cognitive task combination is not enough to elicit cortisol changes; the task must include elements of uncontrollability and social-evaluative threat to strongly and reliably activate this system. Furthermore, the results appear to justify the added effort of utilizing an audience to induce social-evaluative threat (rather than only videotaping the session), as the strongest effects were found when the evaluating individuals were actually present.

Other methodological factors were not associated with cortisol responses across the studies. The method of cortisol assessment (plasma vs. saliva sampling) did not significantly predict the effect sizes. Consistent with past research that has found high correlations between salivary and plasma cortisol responses to acute stressors (Kirschbaum & Hellhammer, 1994), greater changes in cortisol were not associated with one of the assessment methods. The average age or gender of the participants also did not significantly predict the effect sizes. Although some empirical studies that have directly compared cortisol responses in men and women or young and elderly subjects have found small differences between these groups (e.g., Gotthardt et al., 1995; Kirschbaum, Wust, & Hellhammer, 1992), across all of the studies conducted in this area, the gender and average age of the participants do not appear to explain the variability in cortisol responses.

We found that certain methodological features meant to reduce error variance (running all participants at the same time of the day, screening and excluding participants with psychological or physical diseases known to affect the neuroendocrine system) were not associated with cortisol responses. However, this should not be interpreted as evidence that methodological rigor is unimportant; in most cases, the methods section of the articles did not report the detailed information necessary to assess the variety of factors that have been shown to influence cortisol responses. Many studies have documented that behavioral and health factors (e.g., smoking, medication use, physical–psychological disorders; Heim et al., 2000; Kirschbaum et al., 1999; Kirschbaum, Wust, & Hellhammer, 1992) or engaging in certain activities (e.g., consuming caffeine or a meal; al’Absi, Lovallo, McKey, & Pincomb, 1998; Holl, Fehm, Voigt, & Teller, 1984) can affect cortisol responses to psychological stressors. Excluding participants with these health characteristics and providing behavioral restrictions before the laboratory session is clearly the best way to obtain interpretable cortisol results (for discussion, see Kirschbaum & Hellhammer, 1994; Lovallo & Thomas, 2000).

Research syntheses can be useful for computing power analyses to guide decisions about sample sizes to be used in future research. The uncontrollable, social-evaluative tasks generated the largest effect size observed in this review (df = 0.92), and approximately 40 participants would be necessary to have sufficient power (.80) to detect effects. However, tasks with only one of these elements (e.g., only social-evaluative threat or only uncontrollability), associated with an effect size of 0.30–0.35, would require over 200 participants (although fewer would be needed if methodological criteria were optimized). This underscores the need for investigators to determine the specific elements of the stressor to be used in a study before conducting power analyses, as these characteristics are critical factors for determining the appropriate number of participants to enroll in the investigation.

Limitations and Future Directions

Several limitations to this research synthesis warrant comment. We could not test certain alternative explanations for the findings; it is possible that other factors that could not be adequately evaluated in the research synthesis were driving the results. For example, it is possible that some tasks are simply more stressful in some way, and generalized stress may be associated with cortisol responses. However, social-evaluative, uncontrollable tasks were not perceived as more distressing than other types of tasks, suggesting that generalized stress may not explain the results. Although uncontrollability is often confounded with other factors such as effort, difficulty, or unpredictability, several animal and human studies have independently manipulated uncontrollability and these related constructs (e.g., Hess et al., 1983; Peters et al., 1998). These studies have found that uncontrollability is specifically related to acute cortisol responses, reducing the viability of these alternative explanations.

Because the research synthesis was limited to first-time exposure to the laboratory stressor tasks, we could not examine the role of novelty in modulating cortisol responses. However, studies that have participants undergo the same social-evaluative, uncontrollable stressor on multiple days have found significant cortisol elevations on subsequent exposures, indicating that novelty is not the only factor driving cortisol responses (S. Cohen et al., 2000; Kirschbaum, Pruessner et al., 1995; Schommer, Hellhammer, & Kirschbaum, 2003). It is likely that novelty and uncontrollability...
are intertwined, to the extent that when a situation is novel, the outcomes are less certain or controllable. The interactive and unique effects of novelty and other relevant constructs on the HPA system should be addressed in future empirical investigations.

The research synthesis delineated specific factors that elicit cortisol responses in healthy adults, and does not speak to whether the same conditions would provoke cortisol responses in children and adolescents. Several lines of evidence do suggest that the HPA systems of young adults are sensitive to social-evaluative, uncontrollable contexts. Buske-Kirschbaum and colleagues (1997) have shown that a modified version of a public speaking/cognitive task combination (with social-evaluative and uncontrollable elements) does reliably elicit cortisol changes in children and young adults. Gunnar and colleagues have demonstrated, in a series of studies, that socially threatening contexts that are characterized by rejection can increase cortisol levels in children (reviewed in Gunnar & Donzella, 2002). Future research could clarify the role that social evaluation and uncontrollability may play in eliciting cortisol responses in younger populations, and address possible developmental changes in the elicitors of this system.

The research synthesis focused on explaining the cortisol variability between studies in order to address the fundamental question of what specific experimental conditions trigger activation of this system. We were not able to examine individual difference factors and appraisal processes that are clearly important for understanding variability in cortisol changes between individuals. However, a theoretical understanding of the contexts that are reliably associated with cortisol changes can highlight specific classes of individual differences that could be particularly fruitful for future research. For example, social-evaluative conditions may be especially costly to individuals with characteristics that make them sensitive to social rejection, such as social anxiety, rejection sensitivity, or low self-esteem (Downey & Feldman, 1995; Leary, Kowalski, & Campbell, 1988; Nezlek, Kowalski, Leary, Blevins, & Holgate, 1997). These self-related factors have predicted exaggerated cortisol responses to acute stressor tasks (Kirschbaum, Pruessner, et al., 1995; Pruessner et al., 1999; Schmidt et al., 1999; Seeman, Berkman, et al., 1995), but relationships have not emerged between other types of personality variables and cortisol changes (e.g., neuroticism, extraversion, trait anxiety, coping styles; Bossert et al., 1988; Kirschbaum, Bartussek, & Strasburger, 1992; Schommer et al., 2003; van Eck, Nicolson, Berkhof, & Sulon, 1996). It is possible that by matching the context with relevant vulnerability factors (e.g., social-evaluative contexts and individual sensitive to social-evaluation), we can identify those who may be particularly likely to show strong cortisol responses to specific situations.

The meta-analytic findings relate only to short-duration laboratory stressors and cannot address the impact of stressors that can be experienced in the real world; it is unclear the extent to which cortisol responses to acute laboratory stressors mirror those to naturalistic stress. Research in the area of cardiovascular reactivity suggests that the link between reactivity in the laboratory and reactivity in the natural environment may be weak (Gerin, in press). However, research comparing cortisol responses in the laboratory and in naturalistic settings is limited. Laboratory results could provide conservative estimates of the effects of social-evaluative threat and uncontrollability. Experimental conditions are never completely uncontrollable (participants can withdraw at any time), and social-evaluative threat in an experimental setting is not nearly as distressing as it is in real life, where negative evaluation can have profound, long-lasting consequences.

Research coupling experience-sampling methodology with laboratory-based stressor tasks is crucial to determine the generalizability of the experimental findings to real-world contexts. Although individuals are not routinely asked to complete mental arithmetic problems before an audience in daily life, the real world does involve uncontrollable, social-evaluative conditions. For example, academic and professional experiences can provide the potential for failure and negative evaluation by teachers, supervisors, and peers. However, many social-evaluative experiences do not occur within a performance context but, instead, involve negative interpersonal evaluations and/or rejection within ongoing social interactions.

Several lines of research have demonstrated that these negative interpersonal contexts can elicit cortisol responses as well. For example, female participants who were systematically ignored and ostracized by confederates during a laboratory interaction task showed increases in cortisol (Stroud et al., 2000). Children who are rejected by their peers have higher cortisol levels than those who are popular or accepted within their social group (Gunnar & Donzella, 2002). Several laboratory marital interaction studies have found that couples that respond to conflict discussion with hostile behaviors (e.g., criticisms, “put-downs,” and disapprovals) show elevations in cortisol, whereas those with a more positive interaction style do not (Kiecolt-Glaser et al., 1997; Malarkey et al., 1994). These contexts, characterized by explicit rejection and social evaluation, demonstrate that interpersonal interactions that threaten the social self can also trigger activation of the cortisol system; this does not appear to be restricted to performance situations. Relationships that are critical, rejecting, or harassing could create an uncontrollable, evaluative context that could activate the HPA system. These findings point to the importance of evaluating links between specific negative interaction patterns and their physiological correlates in an attempt to better model laboratory tasks after real-world contexts.

Threats to the social self are clearly not the only conditions that elicit cortisol responses. It is relatively easy to threaten the social self in a laboratory context, but threats to other important goals (e.g., loss of a loved one) are much harder to model. Animal studies demonstrate that threats to important goals, such as physical self-preservation, elicit HPA activation. It will be important to determine whether threats to other central goal domains, both within and outside of the laboratory, are also associated with HPA activity in humans.

The research synthesis focused on the HPA system exclusively when, clearly, these specific social conditions elicit a broad range of patterned hormonal and autonomic responses. Future research should extend these findings into other relevant systems, for example, by assessment of catecholamines, cardiovascular responses, other hormonal processes, and the expression of relevant receptors, and should include the pattern of change in these processes over time. The specific neuronal circuits that form the substrates of these coordinated responses should be defined. At the same time, determining the critical environmental signals that most clearly activate these physiological responses to social-evaluative threat, such as particular facial or nonverbal expressions from evaluative
Summary

Overall, despite the inconsistencies and controversies in the literature, we found that acute psychological stressors can elicit cortisol activation. However, not all acute psychological stressors provoke this system; there was a substantial degree of variability in the size of the cortisol effects, depending on the characteristics of the stressor, supporting a stressor-physiology “specificity” perspective. We found strong support for our theoretical model that uncontrollable threats to the social self elicit robust and reliable cortisol responses. Performance tasks characterized by socially-evaluative threat and/or uncontrollability triggered significant elevations in cortisol levels, and the largest increases were found for performance tasks containing both elements. These conditions affected both the overall magnitude of cortisol responses and time to recovery, as well as ACTH responses. Social self-preservation is a key priority across human cultures; threats to this goal may be one important set of eliciting conditions for activating a central physiological system with psychological and health implications.

References

References marked with a single asterisk indicate the studies included in the research synthesis. References marked with double asterisks indicate the studies included in the research synthesis. References marked with double asterisks indicate the studies included in the research synthesis. References marked with double asterisks indicate the studies included in the research synthesis. References marked with double asterisks indicate the studies included in the research synthesis.


hypothalamic secondary amenorrhea. *Fertility and Sterility,* 73, 812–816.


*Green, E. L. (2001). A study of self identified children of alcoholics (COAs) and nonCOAs to determine if they differ in their response to stress and if they differ in their use of alcohol. *Dissertation Abstracts International,* 62(7), 3154B. (UMI No. 3020454)


docrine, and immunological reactivity in response to an acute laboratory stressor. Psychosomatic Medicine, 63, 493–501.


Acute Stressors and Cortisol Responses


Nejtek, V. A. (1999). Salivary cortisol levels indicate emotional arousal-stress magnitude after viewing high and low emotion events: Can salivary cortisol levels also be related to memory for the events? (Doctoral dissertation, University of Texas, Dallas). Dissertation Abstracts International, 60(6), 3007B.


The Yale Interpersonal Stressor (YIPS): Affective, physiological, and behavioral responses to a novel interpersonal rejection paradigm. Annals of Behavioral Medicine, 22, 204–213.


Received November 20, 2002
Revision received August 1, 2003
Accepted August 4, 2003