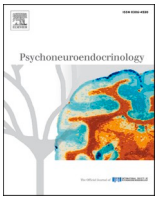
Psychoneuroendocrinology 163 (2024) 106976

Contents lists available at ScienceDirect

Psychoneuroendocrinology

journal homepage: www.elsevier.com/locate/psyneuen



Associations of diurnal cortisol parameters with cortisol stress reactivity and recovery: A systematic review and meta-analysis

Christiane Wesarg-Menzel [a,*], Ruth Marheinecke [a], Janneke Staaks [b], Veronika Engert [a,c,d]

[a] Institute for Psychosocial Medicine, Psychotherapy and Psychooncology, Jena University Hospital, Friedrich-Schiller University, Jena, Germany
[b] University Library, University of Amsterdam, Amsterdam, the Netherlands
[c] Social Stress and Family Health Research Group, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[d] German Center for Mental Health (DZPG), Germany

ARTICLE INFO

Keywords:
Cortisol
HPA axis
Stress
Reactivity
Recovery
TSST (max. 6)

ABSTRACT

Researchers commonly assess the functioning of the hypothalamic-pituitary-adrenal (HPA) axis by measuring natural fluctuations of its end product cortisol throughout the day or in response to a standardized stressor. Although it is conceivable that an individual releasing relatively more cortisol when confronted with a laboratory stressor does the same in everyday life, inconsistencies remain in the literature regarding associations between diurnal cortisol parameters and cortisol stress responses. Hence, the current meta-analysis aggregated findings of 12 studies to examine overall associations of diurnal cortisol parameters (including total output, diurnal slope, and cortisol awakening response [CAR]) with cortisol stress reactivity and recovery in the Trier Social Stress Test (TSST). There were no significant overall associations of total output, slope, or CAR with stress reactivity. Lower total diurnal cortisol output was significantly related to better stress recovery, whereas diurnal slope and CAR were unrelated to stress recovery. Moderation analyses revealed that associations between diurnal cortisol and cortisol stress responses were dependent on the computation method of cortisol parameters, questioning the convergence and validity of commonly employed measures of stress reactivity and recovery. Overall, it seems that we cannot predict characteristics of the diurnal cortisol rhythm from a one-time measure of stress reactivity in a standardized psychosocial laboratory paradigm.

1. Introduction

The hypothalamic-pituitary-adrenal (HPA) axis is a neuroendocrine system extensively studied in the context of stress and disease (DeMorrow, 2018). Researchers commonly assess the functioning of the HPA axis by measuring its end product cortisol in two different contexts: 1) natural fluctuations in cortisol levels throughout the day (termed "diurnal cortisol rhythm"; Ross et al., 2014), and 2) acute levels of cortisol in response to a stressor (Spencer and Deak, 2017; Zänkert et al., 2019). Although reflecting distinct processes of HPA axis regulation, it is reasonable to assume that diurnal cortisol rhythm and cortisol stress response are closely linked, since both are coordinated by the same system. Beyond that, diurnal cortisol rhythm and cortisol stress response are both characterized by individual differences such as sex (Gunn et al., 2016; Juster et al., 2016; Liu et al., 2017), age (Mikneviciute et al., 2023; Shenk et al., 2022) and psychopathology (Fairchild et al., 2008), and are both affected by stressful environments (Raffington et al., 2018). In line with the assumption of an interdependence of these two measures of HPA axis activity, it has previously been suggested that heightened acute stress responses measured in the laboratory generalize to higher levels of diurnal cortisol release (Kidd et al., 2014; Simons et al., 2017). Therefore, an individual releasing relatively more cortisol when confronted with one exemplary standardized stressor is assumed to do the same in everyday life. Yet, only few studies have explicitly investigated associations between measures of diurnal cortisol and cortisol stress response (Kidd et al., 2014; Lucas-Thompson et al., 2018; Simons et al., 2017). As several studies have in fact collected both measures in the same participants, aim of the present meta-analysis was to aggregate the current literature on associations between measures of diurnal cortisol and the cortisol stress response in healthy children, adolescents, and adults, attempting to provide a deeper understanding of mechanisms linking HPA axis (dys-)regulation and disease risk.

* Correspondence to: Institute for Psychosocial Medicine, Psychotherapy and Psychooncology, Jena University Hospital, Friedrich-Schiller University, Stoystr. 3, 07743 Jena, Germany.
E-mail address: christiane.wesarg-menzel@med.uni-jena.de (C. Wesarg-Menzel).

https://doi.org/10.1016/j.psyneuen.2024.106976
Received 6 October 2023; Received in revised form 4 January 2024; Accepted 19 January 2024
Available online 24 January 2024

Cortisol is a glucocorticoid (GC) hormone that is essential for the maintenance of homeostasis and enables the organism to deal with stress (Papadimitriou and Priftis, 2009). It is synthesized in the adrenal cortex and released in reaction to a hormonal cascade initiated in the paraventricular nucleus (PVN) of the hypothalamus. The PVN releases corticotropin-releasing hormone (CRH) and arginine vasopressin (AVP), which trigger the release of adrenocorticotrophic hormone (ACTH) from the anterior pituitary. ACTH in turn triggers synthesis and secretion of GCs from the adrenal cortex (Charmandari et al., 2005). GCs regulate HPA axis activity through negative feedback mechanisms acting at the level of the hippocampus, pituitary gland and the hypothalamus to inhibit further cortisol release (De Kloet et al., 1991; Jacobson and Sapolsky, 1991; Tsigos and Chrousos, 2002).

Cortisol levels follow a diurnal rhythm characterized by a cortisol peak between 30 and 45 min after awakening, followed by a sharp decline over the next hour or two, and a more gradual decline during the remainder of the day (Fries et al., 2009; Kirschbaum and Hellhammer, 1989; Pruessner et al., 1997). Circulating levels of cortisol can easily be determined from saliva, but also from blood and urine (Spencer and Deak, 2017). Stress researchers are commonly interested in three different diurnal cortisol parameters: 1) total output, 2) diurnal slope, and 3) cortisol awakening response (CAR; Hulett et al., 2019). Total cortisol output is estimated for a determined time interval and often calculated as the area under the curve with respect to ground (AUC_G; Pruessner et al., 2003). The diurnal slope describes changes in cortisol levels from morning to evening and can be calculated either as a difference score of the first and last sample of the day, or a regression coefficient, including additional samples collected throughout the day (Adam et al., 2017). The CAR refers to the dynamic change in cortisol from waking to reaching its peak, which occurs within the first hour after waking (Stalder et al., 2016). It is commonly computed as a difference score from waking to 30 min post-waking, or as the AUC with respect to increase (AUC_I) if more than two samples are obtained (Fekedulegn et al., 2007).

On top of the diurnal rhythm, cortisol levels are released in response to acute challenges or stressors, reaching their peak about 20–40 min post-stressor onset, and returning to baseline within 60 min after stressor cessation (Dickerson and Kemeny, 2004). In several studies, the cortisol stress response is divided into a phase of "stress reactivity" that is followed by a phase of "stress recovery" (Nierop et al., 2006). The sampling time point at which individual or average peak cortisol levels occur is often used as a boundary between both phases. To quantify reactivity and recovery, one common procedure is to analyze separate AUC measures for each phase. Regarding cortisol stress reactivity, many authors have computed the AUC_I (e.g., Sandner et al., 2020), which represents reactive increases in cortisol levels in relation to a stressor without regard for baseline cortisol levels (Pruessner et al., 2003). Others have referred to the AUC_G as a measure of stress reactivity (e.g., Gerber et al., 2020; Zorn et al., 2017). However, the AUC_G rather represents total cortisol output (Pruessner et al., 2003) and may hence reflect stress reactivity less accurately than the AUC_I (Khoury et al., 2015). It is also common to analyze difference scores for each phase: for stress reactivity, the difference between peak (maximum) and baseline cortisol levels (*Peak-Baseline*), or maximum and minimum levels (*Max-Min*; e.g., Halbeisen et al., 2023; Khoury et al., 2015); for stress recovery, the difference between baseline and last cortisol value (*BLLV*), or percentage change scores from peak to last value (e.g., Juster et al., 2012; see Table 1). A more recent approach is to employ Piecewise Growth Curve Modeling with Landmark Registration (Lopez-Duran et al., 2014) to model reactivity and recovery slopes relative to the individual cortisol peak (e.g., Degering et al., 2023; Malanchini et al., 2021). Yet, a shared limitation of these approaches is that the extent to which an individual reacts to a stressor may influence subsequent recovery (see e.g., Degering et al., 2023). Moreover, the co-occurrence of processes related to reactivity and recovery hampers disentanglement of both phases (Miller et al., 2018).

Addressing this much debated issue, Miller and colleagues (2018) conducted a systematic evaluation and validation of various parameters

Table 1
Computation of cortisol parameters.

	Parameter	Computation	Interpretation: higher values indicate…
Diurnal cortisol			
Total output (TO)	AUC_G	area under the curve with respect to ground (formula by Pruessner et al., 2003)	higher total cortisol release throughout the day
Slope	Wake-Bed	cortisol at wake-up minus cortisol at bedtime	steeper decline from wake-up to bedtime
	Wake-Late	cortisol at wake-up minus cortisol at late evening	steeper decline from wake-up to late evening
Cortisol awakening response (CAR)	0-30 delta	cortisol at 30 min post wake-up minus cortisol at wake-up	higher increase in cortisol from wake-up to 30 min post wake-up
	Peak-Wake	cortisol at peak concentration within 60 min post wake-up minus cortisol at wake-up	higher increase in cortisol from wake-up to peak cortisol levels
	AUC_I	area under the curve with respect to increase (formula by Pruessner et al., 2003)	higher cortisol release on top of wake-up levels
Stress reactivity			
	AUC_I	area under the curve with respect to increase (formula by Pruessner et al., 2003)	higher cortisol release on top of baseline levels until average peak time point
	AUC_G	area under the curve with respect to ground (formula by Pruessner et al., 2003)	higher cortisol release in response to stress until average peak time point
	MaxMin	peak (maximum) cortisol minus lowest (minimum) value	larger difference in cortisol between peak and lowest concentration of all measurement time points (i.e., lowest concentration can represent baseline, but also any other sample before or after peak)
	Peak-Baseline	peak cortisol minus baseline cortisol	higher cortisol increase from baseline to peak cortisol levels
Stress recovery			
	C_{Min}*	minimal cortisol concentration across all measurement time points (including baseline; see Miller et al., 2018)	weaker recovery
	BLLV	baseline cortisol minus cortisol at last measurement time point	higher cortisol decrease from baseline to last measurement time point (i.e., stronger recovery)
	Percentage Change*	percent change score from 20 min post TSST onset to last measurement time point (formula by Juster et al., 2012)	lower decrease from cortisol at assumed peak to last measurement time point (i.e., weaker recovery)

Note. Exact computation formulas are provided in the analysis script posted on Open Science Framework (https://osf.io/t4dxa/).
*To facilitate interpretability of stress recovery parameters so that higher levels indicate higher / better recovery, effect sizes based on the parameters C_{Min} and Percentage Change were converted * (−1).

commonly used to index stress reactivity and recovery based on a newly developed, physiologically plausible model of cortisol secretion. This pharmacokinetic model included the two major process components of stress-related cortisol secretion (reactivity) and cortisol elimination (recovery), as well as two often neglected components, namely the secretory delay after stress onset, and deviations from the projected steady-state concentration due to stress-unrelated fluctuations of cortisol secretion. After fitting the model to two independent study samples in which participants had completed the TSST, the model was used to examine the correlations of the four components with commonly used indices of stress reactivity and recovery. Based on these analyses, the authors recommended to use *MaxMin* and the minimum cortisol level (C_{Min}) as relatively unadulterated proxy measures of reactivity and recovery, respectively (Miller et al., 2018). Based on these measures, a higher difference between maximum and minimum cortisol levels indicates higher stress reactivity, and lower minimum cortisol levels indicate higher (i.e., better) stress recovery.

Stressor characteristics—next to individual factors (e.g., social anxiety; Crişan et al., 2016) and contextual factors (e.g., the presence of a supportive person; Kirschbaum et al., 1995)—determine the magnitude of the stress response (Dickerson and Kemeny, 2004). The Trier Social Stress Test (TSST; Kirschbaum et al., 1993) is a commonly employed standardized laboratory stress task that has been shown to produce large and reliable HPA axis responses in both healthy and clinical populations (Dickerson and Kemeny, 2004; Zorn et al., 2017). It includes an anticipatory preparation period followed by a task in which participants give a speech and perform a mental arithmetic task in front of an audience showing neutral and reserved behavior (Kirschbaum et al., 1993). The TSST thereby covers all key stressor elements to which the HPA axis sensitively reacts, including novelty, uncertainty, perceived loss of control over the environment, and social-evaluative threat (Dickerson and Kemeny, 2004).

Few studies have investigated associations between diurnal cortisol measures and the cortisol stress response in a systematic way. For instance, it has been observed that higher pre-stress (baseline) cortisol levels were associated with lower cortisol reactivity to the TSST, both in children (Malanchini et al., 2021) and adults (Kudielka et al., 2004). Hence, it is possible that higher diurnal cortisol levels reduce a superimposed stress response, which might indicate impaired ability of the HPA axis to adaptively react to stressors. However, studies in which a diurnal cortisol profile was assessed on one or more additional days do not support this hypothesis, and rather reveal opposite associations. To exemplify, diurnal cortisol output (AUC-derived) was found to positively relate to cortisol responses to cognitive challenges in adults (Kidd et al., 2014) and to social evaluative stress in children (Simons et al., 2017). These findings suggest that variations in cortisol responses may generalize to the amount of cortisol output over the day. Similar to diurnal cortisol output, the diurnal cortisol slope shows only inconsistent associations with cortisol reactivity. Smaller diurnal declines (i.e., less steep slopes) were reported in relation to higher (Simons et al., 2017), but also lower cortisol stress reactivity (Lucas-Thompson et al., 2018), whereas another study observed no association between both measures (Malanchini et al., 2021). With regards to the CAR, several studies revealed no association with stress reactivity (Kidd et al., 2014; Malanchini et al., 2021; Schmidt-Reinwald et al., 1999) or recovery (Malanchini et al., 2021). However, a recent study from our laboratory differentiating between reactivity and recovery found a specific link of a higher CAR with decreased (i.e., more shallow) stress recovery (Degering et al., 2023).

Given these inconsistencies concerning the link between diurnal cortisol measures and the cortisol stress response, aim of the present meta-analysis was to quantify overall effect sizes of associations of total diurnal cortisol output, diurnal slope, and CAR with acute cortisol stress reactivity and recovery. To this end, we focused on stress responses elicited by the TSST, one of the most widely used and potent standardized laboratory stress tests for which an adaptation for children exists (Buske-Kirschbaum et al., 1997). Uncovering associations between basal and reactive HPA axis states may facilitate our understanding of how specific aspects of the diurnal cortisol rhythm can affect an individual's physiological capacity to cope with acute stressors. Vice versa, it may clarify whether we can actually deduce daily HPA axis functioning from a one-time measure of stress reactivity in a standardized psychosocial laboratory paradigm.

2. Methods

2.1. Identification and selection of studies

We identified articles by searching the electronic databases PsycINFO, Web of Science, Embase, and Medline on 28th of November 2022. Keywords related to the diurnal cortisol rhythm were combined with those related to cortisol stress reactivity and recovery (for search strategy, see Supplemental Material). No limits regarding language, publication year or publication type were applied during the search phase. The search results were screened for duplicates with Zotero (for tutorial, see Staaks, 2020).

The first and second author conducted the screening of articles by using the systematic review web application Rayyan (Ouzzani et al., 2016). Ten randomly selected articles served as the training set to establish reliability among coders. In a first stage, both coders screened a random selection of 10% of the reports based on title and abstract to determine eligibility for full-text screening. The agreement between raters was almost perfect (*kappa* = 0.83; 95.45% of agreement). All remaining reports were then split up between the two coders. In a second stage, this procedure was repeated for full-text screening. In this screening stage, agreement between raters was perfect (*kappa* = 1.00; 100% of agreement).

Studies were included if they assessed at least one parameter of the diurnal cortisol rhythm (total diurnal cortisol output, diurnal slope, or CAR) in conjunction with the cortisol stress response to the TSST in healthy children, adolescents, or adults. Diurnal and stress measures were required to be obtained on different days during the same data collection wave. It was further a precondition that total cortisol output was measured with at least three samples, diurnal slope with at least two samples (morning and evening), and CAR with at least two samples (at wake-up and within first hour after wake-up). Measures of diurnal slope and total cortisol output should not contain post wake-up samples pertaining to the CAR (e.g., 30 or 45 min post wake-up) to exclude the state-dependent morning peak (Adam et al., 2017). Stress reactivity was required to be assessed with at least two samples (at baseline and post-stressor-onset), whereas one sample was considered sufficient for stress recovery (if it represented the minimum concentration throughout the TSST; see Miller et al., 2018). We included published articles that were reported in English or German. Studies were excluded if all participants were diagnosed with a psychiatric disorder (e.g., major depression) or a physical disease known to affect HPA axis functioning (e.g., Cushing disease), were pregnant or breast-feeding, or had been exposed to trauma (e.g., child maltreatment or war). If only a subsample of a study met these exclusion criteria, we included data from a control group. If a study met the inclusion criteria but did not report the data necessary to compute an effect size, the corresponding author was contacted. If no data were provided upon request, the study was excluded.

2.2. Data extraction and coding

The first author extracted the data from each study into a coding sheet (see https://osf.io/t4dxa/) and computed the effect sizes. To assess inter-rater reliability, 50% of the studies were randomly selected and scored by the second author. Inter-rater reliability was determined by calculating the percentage of agreement for all study characteristics, Cohen's Kappa for categorical variables and intraclass correlation for continuous variables. The inter-rater agreement for categorical variables

proved to be moderate to almost perfect, with Kappa = 0.70 (83% agreement) for order of diurnal and stress measures, and Kappa = 1.00 (100% agreement) for all other categorical variables (e.g., country and objective monitoring). The inter-rater reliability for continuous variables was excellent, with intraclass correlations ranging from 0.99 (83% agreement) for percentage females to 1.00 (100% agreement) for publication year and number of TSST samples.

2.2.1. Study characteristics

We coded publication year, country of data collection (categorized additionally into North American countries vs. other countries), and data report (raw data or effect sizes provided).

2.2.2. Sample characteristics

We extracted information on sample size, mean age of participants, sex distribution (coded as percentage of females), and ethnicity (coded as percentage of minorities). Other sample characteristics coded were information on participants' socio-economic background (education and income), lifestyle variables (e.g., smoking), physical and mental disorders, as well as menstrual cycle phase and oral contraceptive use for female participants.

2.2.3. Characteristics of cortisol measures

We coded diurnal cortisol parameter (total output, slope, CAR) and its computation method (see Table 1), number and timing of cortisol samples per day, number of sampling days (> 1 day, consecutive or not), number of days between assessment of diurnal cortisol and cortisol stress reactivity, whether participants were forcedly awakened (if CAR assessment took place), and whether objective monitoring of sampling was applied. We further coded duration of the TSST (including speech and arithmetic task), daytime during which the TSST took place, computation of stress reactivity / recovery index (see Table 1), number and timing of samples, TSST responder rate, and characteristics of the committee (i.e., number of committee members, percentage female, mean age).

2.3. Effect size computation

We used the Pearson product-moment correlation coefficient *r* as the common effect size for associations of diurnal cortisol and cortisol stress reactivity / recovery. The majority of effect sizes (72%) was derived from raw data provided by one of the authors of the concerned study. In these data sets, raw cortisol values were log-transformed to correct for significant skew. Then, cortisol samples > 3 SD above the mean were winsorized to reduce outlier effects (Hostinar et al., 2015). Various parameters for diurnal cortisol and cortisol stress reactivity / recovery were computed based on an SPSS script (see https://osf.io/t4dxa/). In line with previous studies, stress reactivity and recovery were separated from one another by the sample's average cortisol peak. For the remaining 28% of studies, correlations were directly provided by one of the authors or were reported in the manuscript. A positive correlation coefficient reflected that higher stress reactivity (i.e., higher cortisol increase) or higher / better stress recovery (i.e., stronger decrease from baseline to last cortisol sample; lower C_{Min}) were associated with higher levels of diurnal cortisol parameters (i.e., higher total diurnal cortisol output, steeper cortisol slopes, and / or higher CAR). To this end, we converted effect sizes based on the recovery parameters C_{Min} and Percentage Change from peak to last measurement time point by multiplying them with (−1). This procedure facilitated interpretation of stress recovery parameters such that higher levels indicate higher / better recovery (see Table 1). For meta-analyses, we transformed Pearson correlation coefficients into Fisher's *Z* scores, as the latter are normally distributed (Lipsey and Wilson, 2001). To facilitate interpretability, we converted Fisher's *Z* scores back into Pearson correlations after performing statistical analyses.

2.4. Statistical analysis

We conducted six meta-analyses to assess overall associations of three diurnal cortisol parameters (i.e., total output, slope, CAR) with cortisol stress reactivity and recovery in the TSST. Given that several studies reported on various cortisol parameters and employed multiple computation methods to estimate these, it was possible to extract multiple effect sizes from all studies. To deal with effect size dependency, we applied a multi-level approach (Assink and Wibbelink, 2016; Cheung, 2015; Van den Noortgate et al., 2013). Effect sizes were pooled in a three-level random effects model, in which three levels of variance were modeled: variation in effect sizes due to random sampling of effect sizes (Level 1), variation in effect sizes due to differences within a single study (Level 2), and variation in effect sizes between different studies (Level 3; Assink and Wibbelink, 2016). We assessed the overall strength of association between diurnal cortisol parameters and both stress reactivity and recovery by building separate three-level meta-analytic models without predictors (i.e., intercept-only models), in which the estimated intercept values represented the corresponding associations.

To determine heterogeneity in effect sizes, we assessed the distribution of the variance over the three levels of the meta-analytic model (i.e., sampling variance at level 1, within-study variance at level 2, and between-study variance at level 3; Assink and Wibbelink, 2016). We applied the rule by Hunter and Schmidt (1990) who state that heterogeneity can be considered as substantial if less than 75% of the total amount of variance can be attributed to sampling variance. In case of significant heterogeneity, moderator analyses were conducted. We created dummy variables for each discrete variable, and mean centered all continuous variables. Each moderator was then added as a covariate to a separate three-level meta-analytic model and compared via omnibus tests to the meta-analytic model without predictors (Assink and Wibbelink, 2016). Categorical moderator analyses were restricted to cases where each category included at least five studies. This prevented testing for a moderating role of TSST daytime (with most studies conducting the TSSTs within a time period spanning afternoon and evening), data report (raw data vs. effect sizes provided), and country of data collection (the latter concerned only the association of total cortisol output with stress reactivity and recovery).

All analyses were conducted in R (version 4.2.3; R Core Team, 2023) using the *metafor* package (Viechtbauer, 2010) and syntax provided by Assink and Wibbelink (2016). A significance level of $p < 0.05$ was used in all analyses. In line with guidelines provided by Fernández-Castilla et al. (2020), we created summary forest plots, which account for the existence of multiple effect sizes within primary studies. The dataset and the R syntax employed for the analyses are provided on the Open Science Framework (OSF; https://osf.io/t4dxa/).

3. Results

3.1. Study selection and characteristics

Fig. 1 shows the PRISMA flow diagram (Page et al., 2021) for the identification of articles. The search process yielded 962 potentially relevant records, from which 436 remained after deduplication. After title and abstract screening, 116 reports remained for full-text screening, from which 101 were retrieved and assessed for eligibility. Based on pre-determined exclusion criteria, 43 reports were excluded. In addition, 44 potentially eligible reports were excluded because authors did not provide effect sizes or raw data upon request. Overall, 14 studies were included from the systematic search, and one study was added from personal knowledge. Hence, the final meta-analytic data set included 15 articles reporting on 12 independent studies (k) comprising 271 effect sizes (#ES). The included studies were published between 2006 and 2023 (median = 2017). Studies were predominantly conducted in Canada ($n = 4$), followed by Germany ($n = 3$), the U.S. ($n = 2$), and Japan, The Netherlands, and Switzerland ($n = 1$ each). The overall

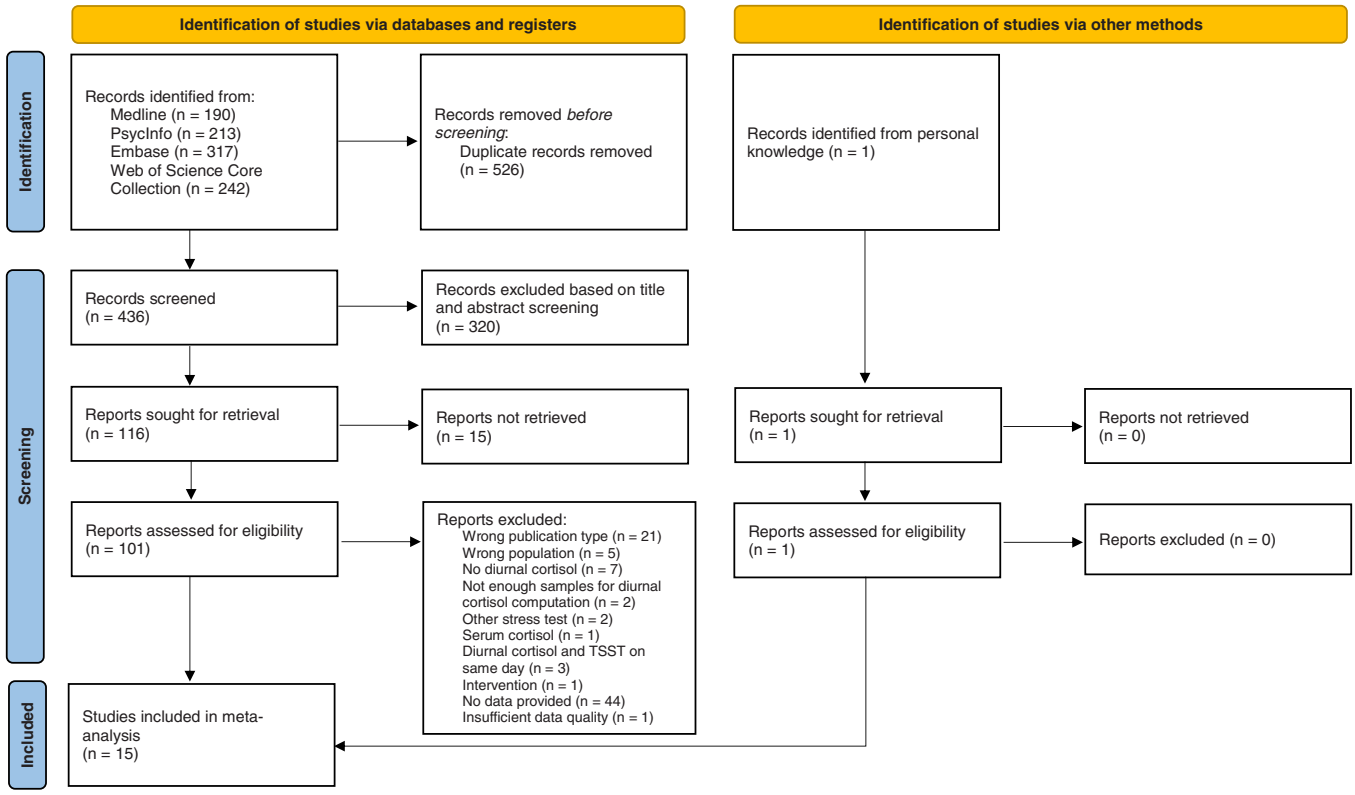


Fig. 1. *PRISMA flow diagram.*

sample included $N = 1235$ unique participants, with study sample sizes ranging from $N = 21$ (Wolfram et al., 2013) to $N = 400$ participants (Malanchini et al., 2021). Participants' mean age was 26.32 years ($SD = 13.84$ years). The mean count of extracted effect sizes per study was 22.58 ($SD = 9.44$). With regard to diurnal cortisol parameters, the majority of effect sizes related to the CAR (n = 140), followed by diurnal slope (n = 70), and total diurnal output (n = 61). On average, diurnal cortisol samples were collected on 1.9 days ($SD = 0.8$). Total diurnal output was assessed with $M = 4.3$ samples ($SD = 1.5$), and the CAR with $M = 2.7$ samples ($SD = 1.0$). In four studies, diurnal sampling times were monitored objectively, and in one study, participants were forcedly awakened. With regard to the stress response, 155 effect sizes pertained to stress reactivity and 116 to stress recovery, with seven TSST cortisol samples being taken on average ($SD = 2$). In most studies, the TSST was conducted within a time period spanning afternoon and evening ($n = 5$), or afternoon only ($n = 3$). Peak cortisol levels occurred at a median of 20 min post-stressor onset. On average, 50.0% ($SD = 23.6$) of the participants responded to the TSST with a baseline-to-peak cortisol release of > 1.5 nmol/l, demonstrating successful stress induction (Miller et al., 2013). Table 2 summarizes study characteristics of included studies (more detailed coding available on OSF; https://osf.io/t4dxa/).

3.2. Associations between diurnal cortisol parameters and stress reactivity

No significant overall associations were found between diurnal cortisol parameters (including total output, slope, and CAR) and stress reactivity (see Table 3 and Figs. S1–3). In each meta-analysis, less than 75% of the total amount of variance could be attributed to random sampling variance, indicating substantial heterogeneity in effect sizes (Hunter & Schmidt, 1990). Hence, we proceeded with moderator analyses (see Table 4).

Reactivity computation method moderated the link between total diurnal cortisol output and stress reactivity. Total diurnal cortisol output was positively related to stress reactivity when reactivity was computed as AUC_G ($r = .244$, $p < .001$), but not when reactivity was computed as AUC_I ($r = -.056$, $p = .274$), Peak-Baseline ($r = -.096$, $p = .078$), or MaxMin ($r = -.075$, $p = .143$). Reactivity computation method also moderated the association between diurnal slope and stress reactivity. Less steep diurnal slopes were related to higher stress reactivity when reactivity was computed as AUC_G ($r = -.127$, $p = .016$), but not when reactivity was computed as AUC_I ($r = .031$, $p = .542$), Peak-Baseline ($r = .048$, $p = .359$), or MaxMin ($r = .050$, $p = .331$). No significant moderator was detected in the meta-analysis on the association between the CAR and stress reactivity.

3.3. Associations between diurnal cortisol parameters and stress recovery

Overall associations between diurnal cortisol parameters and cortisol stress recovery are presented in Table 3. Lower total diurnal cortisol output was significantly related to higher (i.e., better) stress recovery ($r = -.101$; $p = .017$; see Fig. S4). Diurnal slope and CAR were not significantly related to stress recovery (see Figs. S5 and S6).

Given substantial heterogeneity in effect sizes within each meta-analysis (indicated by the fact that less than 75% of the total amount of variance could be attributed to random sampling variance), we proceeded with moderator analyses (see Table 5). Recovery computation method moderated the association between total diurnal cortisol output and cortisol stress recovery. Lower total cortisol output was related to better stress recovery when recovery was computed as C_{Min} ($r = -.253$, $p < .001$), but not when recovery was computed as BLLV ($r = -.034$, $p = .550$) or Percentage Change ($r = .004$, $p = .944$). Recovery computation method further moderated the association between diurnal slope and cortisol stress recovery. In specific, steeper diurnal slopes were related to better stress recovery when recovery was computed as C_{Min} ($r = .141$, $p = .004$), but not when recovery was computed as BLLV ($r = -.007$, $p = .877$), or Percentage Change ($r = .045$, $p = .342$). Publication year, number of participants, and number of TSST samples moderated the association between the CAR and stress recovery. In

Table 2
Study characteristics of included studies.

Study	Country	Sample description	N (% female)	Mean age (SD) in years	Diurnal cortisol parameter and assessment times[a]	Stress response assessment times[b]
Degering et al. (2023)	Germany	healthy adults from general population	129 (54%)	40.4 (9.0)	TO: wake, + 240, + 360, + 480, + 600 Slope: wake, evening CAR: wake, + 30, + 60	− 55, + 10, + 20, + 30, + 40, + 55
Ellenbogen et al. (2006)	Canada	adolescents of parents with no mental disorder[f]	25 (48%)	16.6 (2.1)	TO: wake, 3 PM, 8PM, bed Slope: wake, bed CAR: wake, + 30, + 60	− 40, − 25, − 10, 0, + 5, + 10, + 20, + 30, + 40, + 55 [g]
Evans et al. (2013)	The Nether-lands	children and adolescents from general population[f]	71 (61%)	16.1 (2.5)	TO: wake, 12PM, 8PM Slope: wake, 8PM CAR: wake, + 30	0, + 4, + 18, + 23, + 38, + 53
Juster et al. (2011)	Canada	well-educated full-time workers	27 (63%)	45.4 (2.1)	TO: wake, 2PM, 4PM, bed Slope: wake, bed CAR: wake, + 30	− 50, − 30, − 20, − 10, 0, + 10, + 20, + 30, + 40, + 50 [g]
Juster et al. (2016)[c]	Canada	employees of psychiatric hospitals	188 (71%)	40.4 (0.9)	TO: wake, 2 PM, 4 PM, bed Slope: wake, bed CAR: wake, + 30	− 10, 0, + 10, + 20, + 30, + 40
Lucas-Thompson et al. (2018)	U.S.	children and adolescents from general population	102 (53%)	12.8 (2.2)	TO: wake, + 240, + 360, + 480, + 600 Slope: wake, bed CAR: wake, + 30	0, + 15, + 25, + 35, + 45
Lupien et al. (2022)	Canada	healthy adults from general population	108 (56%)	32.5 (9.3)	CAR: wake, + 30 TO: wake, 2 PM, 4 PM, bed Slope: wake, bed	− 30, − 20, − 10, 0, + 10, + 20, + 30, + 40
Malanchini et al. (2021)[d]	U.S.	children with heterogeneous SES	400 (52%)	10.8 (1.8)	Slope: wake, bed CAR: wake, + 30	− 30, + 20, + 40, + 60
Raffington et al. (2018)	Germany	children with heterogeneous SES	97 (46%)	7.2 (0.5)	Slope: wake, eve CAR: wake, + 15, + 30	− 10, 0, + 10, + 20, + 30, + 40, + 50, + 60 [g]
Wirtz et al. (2007)[e]	Switzerland	healthy men from general population	41 (0%)	42.5 (2.0)	TO: 8AM, 11AM, 4PM, 8PM Slope: wake, 8PM CAR: wake, + 15, + 30, + 45, + 60	− 10, + 10, + 20, + 30, + 40, + 50, + 60, + 70 [g]
Wolfram et al. (2013)	Germany	healthy student teachers	21 (43%)	30.7 (6.3)	CAR: wake, + 30, + 45, + 60	0, + 13, + 23, + 33, + 43, + 58, + 73, + 103 [g]
Yamanaka et al. (2019)	Japan	healthy young adults recruited at university	26 (26%)	20.5 (1.9)	TO: wake, + 120, + 240, + 360, + 480, + 600, + 720, + 840 Slope: wake, + 840	− 35, + 10, + 20, + 30, + 40 [g]

Note. CAR = cortisol awakening response, SES = socio-economic status, TO = total output
[a] + *number* refers to minutes after wake.
[b] − / + *number* refers to minutes relative to stressor onset (coded as 0 = baseline; in most cases, sample immediately before speech task).
[c] Same data set as in Juster et al. (2016) is included in Kerr et al. (2020).
[d] Same data set as in Malanchini et al. (2021) is included in Raffington et al. (2022).
[e] Same data set as in Wirtz et al. (2007) is included in Wirtz et al. (2008).
[f] Control group of study included in meta-analysis.
[g] Sampling times reported in paper were converted to enable comparability with other studies, following the way of reporting described under [b].

Table 3
Estimated pooled correlations (Fisher's Z and Pearson's r) for the relationships between diurnal cortisol parameters with cortisol stress reactivity and recovery.

Association	*#k*	#ES	*N*	Z_r (*SE*)	*r* (*SE*)	95% CI (*r*)	*t* (*df*)	*p*
Diurnal cortisol – Stress reactivity								
Total output – Stress reactivity	9	35	702	.007 (.036)	.007 (.036)	[− .066,.081]	0.20 (34)	.844
Slope – Stress reactivity	11	40	1199	.012 (.038)	.012 (.038)	[− .065,.088]	0.31 (39)	.762
CAR – Stress reactivity	11	80	1209	.011 (.049)	.011 (.049)	[− .086,.108]	0.23 (79)	.816
Diurnal cortisol – Stress recovery								
Total output – Stress recovery	9	26	702	-.101 (.039)	-.101 (.039)	[− .182, − .019]	-2.55 (25)	.017 *
Slope – Stress recovery	11	30	1199	.063 (.031)	.063 (.031)	[− .000,.126]	2.04 (29)	.051
CAR – Stress recovery	11	60	1209	-.022 (.025)	-.022 (.025)	[− .073,.029]	-0.87 (59)	.388

Note. *#k* = number of studies; #ES = number of effect sizes; *N* = total number of unique participants; Z_r = Fisher's Z correlation; *SE* = standard error; *r* = Pearson's correlation coefficient *r*; 95% CI = 95% confidence intervals of Pearson's *r*; *df* = degrees of freedom
*$p < .05$

detail, the strength of the overall association between the CAR and stress recovery became less negative as studies were published more recently, included more participants, and less TSST samples.

3.4. Exploratory analyses

To facilitate interpretability of meta-analytic associations of diurnal cortisol parameters with stress reactivity and recovery, we additionally computed pooled correlations for the relationships of 1) diurnal cortisol parameters, and 2) cortisol stress reactivity and recovery. Steeper diurnal slopes were significantly related to lower CAR ($r = -.371$, $p < .001$) and lower total cortisol output ($r = -.435$, $p < .001$). CAR and total cortisol output were not significantly associated ($r = -.024$, $p = .729$). Most stress reactivity indices were significantly positively correlated, except for AUC_G and MaxMin, which showed no significant association (see Table S1). Stress recovery indices were not significantly

Table 4
Diurnal cortisol parameters and cortisol stress reactivity: Estimated results (Fisher's Z, regression coefficients, omnibus-test) for continuous and categorical moderator variables.

Moderator	#k[a]	#ES	Z_r (*SE*)	β_1 (*SE*)	*F* (*df*1, *df*2)[b]	*p*
Total cortisol output and stress reactivity						
Study characteristics						
Year of publication[c]	9	35	-.001 (.037)	.007 (.007)	$F(1,33) = 1.08$	.306
Sample Characteristics						
Number of participants[c]	9	35	.009 (.040)	-.000 (.001)	$F(1,33) = 0.03$	.862
Age (in years)[c]	9	35	.009 (.036)	-.004 (.003)	$F(1,33) = 1.35$	.253
Sex (% females)[c]	9	35	-.000 (.036)	.002 (.002)	$F(1,33) = 1.66$	.207
Ethnicity (% Minorities)[c]	4	15	.072 (.060)	.001 (.004)	$F(1,13) = .02$	.890
Characteristics of cortisol measures						
Number of total output samples[c]	9	35	.007 (.038)	-.003 (.030)	$F(1,33) = 0.01$	.925
Number of total output sampling days[c]	9	35	.006 (.038)	.034 (.086)	$F(1,33) = 0.16$	.692
Number of TSST samples[c]	9	35	.009 (.038)	.006 (.022)	$F(1,33) = 0.07$	.791
TSST responder rate[c, d]	9	35	.006 (.035)	-.002 (.002)	$F(1,33) = 2.19$	.148
Reactivity computation method:					$F(3,31) = 14.81$	< .001 * **
AUC_G (RC)	9	9	.249 (.050)* **			
AUC_I	9	9	-.056 (.050)	-.305 (.060)* *		
Peak-Baseline	8	8	-.096 (.053)	-.345 (.061)* **		
MaxMin	9	9	-.075 (.050)	-.325 (.059)* **		
Diurnal cortisol slope and stress reactivity						
Study characteristics						
Year of publication[c]	11	40	.011 (.040)	.001 (.008)	$F(1,38) = 0.02$	.899
Country of data collection:					$F(1,38) = 0.51$	.481
U.S. or Canada (North-America; RC)	6	21	.036 (.052)			
Other countries outside North-America	5	19	-.019 (.057)	-.055 (.077)		
Sample Characteristics						
Number of participants[c]	11	40	-.000 (.039)	.000 (.000)	$F(1,38) = 1.22$	.277
Age (in years)[c]	11	40	.014 (.038)	.002 (.003)	$F(1,38) = 0.71$	.404
Sex (% females)[c]	11	40	.017 (.038)	-.002 (.002)	$F(1,38) = 1.30$	.262
Ethnicity (% Minorities)[c]	6	20	-.044 (.045)	.006 (.003)*	$F(1,18) = 4.90$	.040 *
Characteristics of cortisol measures						
Number of slope sampling days[c]	11	40	.005 (.039)	.043 (.050)	$F(1,38) = 0.74$	.396
Slope computation:					$F(1,38) = 0.51$	.481
Wake-Bedtime (RC)	6	21	.036 (.052)			
Wake-Evening	5	19	-.019 (.057)	-.055 (.077)		
Number of TSST samples[c]	11	40	.000 (.036)	-.028 (.020)	$F(1,38) = 1.96$	.170
TSST responder rate[c, d]	10	39	.001 (.040)	.002 (.002)	$F(1,37) = 1.20$	.280
Reactivity computation method:					$F(3,35) = 5.14$	.005 * *
AUC_G	10	10	-.128 (.051)*			
AUC_I	10	10	.031 (.051)	.159 (.053)* *		
Peak-Baseline	9	9	.048 (.052)	.176 (.055)* *		
MaxMin	10	10	.050 (.050)	.177 (.053)* *		
Cortisol awakening response and stress reactivity						
Study characteristics						
Year of publication[c]	11	80	.016 (.053)	-.003 (.009)	$F(1,78) = 0.14$	.713
Country of data collection:					$F(1,78) = 0.24$	.627
U.S. or Canada (North-America; RC)	6	29	.034 (.070)			
Other countries outside North-America	5	51	-.016 (.075)	-.050 (.102)		
Sample Characteristics						
Number of participants[c]	11	80	.004 (.054)	.000 (.001)	$F(1,78) = 0.17$	.678
Age (in years)[c]	11	80	.005 (.047)	-.005 (.003)	$F(1,78) = 2.14$	.148
Sex (% females)[c]	11	80	.009 (.054)	.000 (.003)	$F(1,78) = 0.01$	.913
Ethnicity (% Minorities)[c]	6	28	.089 (.033)*	-.003 (.002)	$F(1,26) = 2.87$	.102
Characteristics of cortisol measures						
Number of CAR samples[c]	11	80	.012 (.049)	.007 (.024)	$F(1,78) = 0.07$	.788
Number of CAR sampling days[c]	11	80	.014 (.053)	-.017 (.069)	$F(1,78) = 0.06$	.807
CAR computation method:					$F(2,76) = 0.20$	.821
AUC_I (RC)	5	20	.024 (.063)			
0-30 delta	10	39	.003 (.056)	-.022 (.041)		
Peak-Wake	5	20	.001 (.063)	-.023 (.042)		
Number of TSST samples[c]	11	80	.013 (.054)	.003 (.028)	$F(1,78) = 0.01$	.922
TSST responder rate[c, d]	10	79	.008 (.061)	.000 (.003)	$F(1,77) = 0.01$	.930
Reactivity computation method:					$F(3,75) = 0.25$	.865
AUC_G (RC)	10	20	.016 (.059)	-.008 (.039)		
AUC_I	10	20	.008 (.059)	-.001 (.039)		
Peak-Baseline	9	19	.015 (.060)			
MaxMin	10	20	-.013 (.059)	-.029 (.039)		

Note. #k = number of studies; #ES = number of effect sizes; Z_r = Fisher's Z correlation; *SE* = standard error; β_1 = estimated regression coefficient; *df* = degrees of freedom; AUC_I = area under the curve with respect to increase; AUC_G = area under the curve with respect to ground; CAR = cortisol awakening response; MaxMin = maximum value minus minimum value; Peak-Baseline = peak value minus baseline value; RC = reference category.
[a] For categorical moderators, a study can be represented in more than one moderator category if the effect sizes reported relate to different moderator categories. Thus, the sum of the number of studies (k) included in distinct categories belonging to the same moderator can exceed 12 (the total number of unique studies); [b] Omnibus test of all regression coefficients in the model; [c] continuous variable; [d] defined as percentage of participants showing a cortisol increase > 1.5 nmol/l from baseline to individual peak levels throughout the TSST.
* $p < .05$; * * $p < .01$; * ** $p < .001$

associated with each other (see Table S1). Overall, higher cortisol stress reactivity was related to lower (i.e., worse) recovery ($r = -.251$, $p < .001$). This association was dependent on reactivity and recovery computation method: Higher reactivity computed as AUC_I or Peak-Baseline was significantly related to lower recovery computed as BLLV. Further, higher reactivity computed as AUC_G was significantly related to lower recovery computed as C_{Min}. None of the stress reactivity measures were significantly related to the recovery index Percentage Change, and none of the stress recovery measures were significantly related to the reactivity index MaxMin (see Table S1).

4. Discussion

The present meta-analysis aimed to uncover associations between common measures of reactive and diurnal HPA axis activity to address the question whether heightened acute stress responses measured in the laboratory generalize to higher levels of cortisol release in daily life. To this end, we included 12 studies in which participants underwent the TSST to assess stress reactivity and recovery, and collected diurnal cortisol samples on one or more additional days. There were no significant overall associations of total diurnal cortisol output, slope, or CAR with stress reactivity. Yet, moderation analyses revealed that higher total diurnal cortisol output and less steep slopes were related to higher stress reactivity when reactivity was computed as AUC_G. Lower total diurnal cortisol output was significantly related to higher (i.e., better) stress recovery, particularly when recovery was computed as C_{Min}. Further, steeper diurnal slopes were only associated with better stress recovery when C_{Min} was used as a recovery index. The CAR was unrelated to stress recovery. Exploratory analyses revealed that higher stress reactivity was moderately related to lower (i.e., worse) stress recovery.

Our finding that total diurnal cortisol output was only related to stress reactivity when reactivity was computed as AUC_G may trace back to the fact that an AUC_G measure of stress reactivity emphasizes the baseline component of cortisol release, as absolute levels of cortisol are incorporated into the computation (see Pruessner et al., 2003). The acute stressor-specific cortisol release is rather neglected in the AUC_G computation. Reactivity measures that are relatively free of diurnal cortisol release such as the AUC_I or MaxMin, on the other hand, did not relate to total diurnal cortisol output. This finding is in line with a study by Simons et al. (2017) in which cortisol responses to the Children's Reactions to Evaluation Stress Test (de Weerth et al., 2013) were examined. The authors found that total diurnal cortisol output was significantly positively related to what they referred to as "total stress cortisol" calculated as AUC_G, but not to stress reactivity computed as standardized residual of a regression predicting peak response cortisol from baseline (Simons et al., 2017). Taking these findings together, we thus conclude that heightened acute stress responses measured in the laboratory do not generalize to higher levels of cortisol release in daily life. Further, in line with previous calls (Khoury et al., 2015), we advise to refrain from labeling AUC_G stress measures as "stress reactivity", given their conflation with diurnal cortisol levels. Still, it is valuable to complement analyses of stress reactivity indices such as the MaxMin (as recommended by Miller et al., 2018) or AUC_I with AUC_G, bearing in mind that they reveal different information (Pruessner et al., 2003).

Lower total diurnal cortisol output was significantly related to better stress recovery. However, moderation analyses revealed that this association was only evident when recovery was computed as the minimum cortisol level throughout the entire TSST procedure (C_{Min}), but not based on other computation methods such as the percentage change score from average peak to last value, or the difference between baseline and last value (BLLV). If C_{Min} is indeed the most accurate index of recovery as proposed by Miller et al. (2018), our finding would suggest that those with lower daily cortisol release are better at recovering from stress. Yet this would question the validity of percentage change scores and BLLV as indicators of recovery, which were not related to C_{Min} and total diurnal cortisol output. Alternatively, one could argue that total diurnal cortisol output does not relate to stress recovery, and question the validity of C_{Min} as a recovery index. In this case, a similar conflation problem may be involved in the observed association between total diurnal cortisol output and stress recovery as was discussed for the AUC_G measure of stress reactivity: Participants with lower total diurnal cortisol output per se may automatically show lower minimum cortisol values during the TSST when compared to those with higher total diurnal cortisol output.

We found no overall associations of diurnal slope with stress reactivity and recovery. However, moderation analyses revealed that steeper diurnal slopes were associated with better stress recovery when C_{Min} was used as a recovery index, but not when percentage change score or BLLV were employed. Assuming C_{Min} as a valid indicator of stress recovery, this finding could suggest that individuals who demonstrate steeper declines in cortisol throughout the day are better able to recover from stress. This would be in line with evidence that both flat diurnal curves and poor stress recovery are associated with poor physical and mental health (Adam et al., 2017; Schoorl et al., 2016), suggesting a shared mechanism within diurnal and reactive HPA axis regulation involved in disease risk. Yet it is also plausible to argue that it is exactly the steeper daily cortisol decline that leads to lower minimum cortisol levels during the TSST, particularly if the TSST is conducted later in the day.

Our meta-analysis revealed that the CAR was unrelated to stress reactivity and recovery, and none of the moderators tested affected this association. In accordance with the present results, Abelson et al. (2023) found no link between the CAR assessed with three salivary samples over each of six days and cortisol stress responses assessed with blood samples drawn at nine times throughout the TSST in 140 healthy participants. Taken together, these findings support conclusions drawn in expert consensus guidelines on CAR assessment that the CAR is distinct from stress reactivity (Stalder et al., 2016). A lack of association between the CAR and cortisol reactivity could also relate to the great extent to which both are determined by different situational factors. The CAR is influenced by waking time, prior day experiences, and anticipation of challenge during the day ahead (Law et al., 2013), whereas cortisol stress reactivity depends on anticipatory and momentary threat appraisal (Gaab et al., 2005; Juster et al., 2012; Schlotz et al., 2011).

Overall, it seems that we cannot predict characteristics of the diurnal cortisol rhythm (including total output, slope and CAR) from a one-time measure of stress reactivity in a standardized psychosocial laboratory paradigm (and vice versa). However, the lack of an association between cortisol stress reactivity and diurnal HPA axis functioning may draw back to the fact that the TSST represents an extremely stressful condition that is rarely encountered in everyday life. More moderately stressful challenge tasks might have higher ecological validity due to closer correspondence with daily life demands. Yet, our results demonstrate that a better recovery from stress as indexed by lower C_{Min} relates to lower total cortisol output over the day and steeper diurnal slopes, which may all represent adaptive patterns of HPA axis functioning.

4.1. Limitations

Important limitations of this work need to be mentioned. At the level of primary studies, it is notable that diurnal cortisol sampling occurred on average on 1.9 days. Methodological studies have indicated that at least three days are necessary to obtain reliable intra-individual measures of total cortisol output (Segerstrom et al., 2014), 5–8 days for diurnal slope (Segerstrom et al., 2014), and six days for CAR (Hellhammer et al., 2007). On the other hand, participant burden and financial costs often obstruct implementation of such extensive sampling protocols. With regards to the CAR, expert consensus guidelines advise for adult studies to take at least three samples (wake-up, +30 min, +45 min) to capture cortisol peak concentrations (Stalder et al., 2016). As the majority of adult studies included in this meta-analysis was published prior to these guidelines, it is perhaps not surprising that only one out of six adult studies included these recommended CAR sampling

Table 5
Diurnal cortisol parameters and cortisol stress recovery: Estimated results (Fisher's Z, regression coefficients, omnibus-test) for continuous and categorical moderator variables.

Moderator	#k[a]	#ES	Z_r (*SE*)	β_1 (*SE*)	*F* (*df*1, *df*2)[b]	*p*
Total cortisol output and stress recovery						
Study characteristics						
Year of publication[c]	9	26	-.097 (.041)*	-.003 (.007)	$F(1, 24) = 0.18$	.674
Sample Characteristics						
Number of participants[c]	9	26	-.104 (.042)*	.000 (.001)	$F(1, 24) = 0.07$	.799
Age (in years)[c]	9	26	-.103 (.040)*	.004 (.003)	$F(1, 24) = 1.26$	.272
Sex (% females)[c]	9	26	-.099 (.041)*	-.001 (.002)	$F(1, 24) = 0.06$	.807
Ethnicity (% Minorities)[c]	4	11	-.125 (.094)	.001 (.006)	$F(1, 9) = 0.01$	.906
Characteristics of cortisol measures						
Number of total output samples[c]	9	26	-.101 (.040)*	-.001 (.033)	$F(1, 24) = 0.00$	.973
Number of total output sampling days[c]	9	26	-.106 (.040)*	.079 (.092)	$F(1, 24) = 0.74$	.399
Number of TSST samples[c]	9	26	-.098 (.041)*	.008 (.024)	$F(1, 24) = 0.11$	.747
TSST responder rate[c, d]	9	26	-.100 (.040)*	.001 (.002)	$F(1, 24) = 0.48$	.494
Recovery computation method:					$F(2, 23) = 6.37$	.006 * *
C_{Min} (RC)	9	9	-.259 (.056)* **			
BLLV	9	9	-.034 (.056)	.225 (.079)* *		
Percentage Change	8	8	.004 (.059)	.263 (.081)* *		
Diurnal cortisol slope and stress recovery						
Study characteristics						
Year of publication[c]	11	30	.066 (.034)	-.001 (.006)	$F(1, 28) = 0.03$	.866
Country of data collection:					$F(1, 28) = 1.92$	.177
U.S. or Canada (North-America; RC)	6	16	.026 (.039)			
Other countries outside North-America	5	14	.108 (.045)*	.082 (.059)		
Sample Characteristics						
Number of participants[c]	11	30	.069 (.034)	-.000 (.000)	$F(1, 28) = 0.20$	.658
Age (in years)[c]	11	30	.056 (.027)	-.004 (.002)	$F(1, 28) = 3.54$	.071
Sex (% females)[c]	11	30	.065 (.033)	-.000 (.002)	$F(1, 28) = 0.01$	.928
Ethnicity (% Minorities)[c]	6	15	.110 (.039)	-.003 (.002)	$F(1, 13) = 2.06$	.175
Characteristics of cortisol measures						
Number of slope sampling days	11	30	.068 (.034)	-.019 (.042)	$F(1, 28) = 0.20$	.662
Slope computation:					$F(1, 28) = 1.92$	.177
Wake-Bedtime (RC)	6	16	.026 (.039)			
Wake-Evening	5	14	.108 (.045)	.082 (.059)		
Number of TSST samples	11	30	.065 (.034)	.001 (.019)	$F(1,28) = 0.01$	.943
TSST responder rate[d]	10	29	.060 (.035)	-.002 (.002)	$F(1,27) = 1.34$	.257
Recovery computation method:					$F(2, 26) = 4.34$	.024 *
C_{Min} (RC)	10	10	.142 (.045)* *			
BLLV	10	10	-.007 (.045)	-.149 (.051)* *		
Percentage Change	9	9	.045 (.047)	-.096 (.053)		
Cortisol awakening response and stress recovery						
Study characteristics						
Year of publication[c]	11	60	-.043 (.019)*	.009 (.003)* *	$F(1, 58) = 7.66$	.008 * *
Country of data collection:					$F(1, 58) = 0.18$	.677
U.S. or Canada (North-America; RC)	6	22	-.034 (.038)			
Other countries outside North-America	5	38	-.011 (.039)	.023 (.055)		
Sample Characteristics						
Number of participants[c]	11	60	-.039 (.022)	.000 (.000)*	$F(1, 58) = 4.52$	.038 *
Age (in years)[c]	11	60	-.022 (.028)	.001 (.002)	$F(1, 58) = 0.06$	.811
Sex (% females)[c]	11	60	-.031 (.025)	.002 (.001)	$F(1, 58) = 1.99$	.164
Ethnicity (% Minorities)[c]	6	21	-.007 (.033)	-.001 (.002)	$F(1, 19) = 0.09$	.768
Characteristics of cortisol measures						
Number of CAR samples[c]	11	60	-.023 (.025)	-.008 (.027)	$F(1, 58) = 0.10$	.759
Number of CAR sampling days[c]	11	60	-.029 (.025)	.037 (.033)	$F(1, 58) = 1.25$	.268
CAR computation method:					$F(2, 56) = 0.325$	.724
AUC_I (RC)	5	15	-.041 (.044)			
0-30 delta	10	29	-.020 (.030)	.021 (.046)		
Peak-Wake	5	15	-.056 (.044)	-.015 (.050)		
Number of TSST samples[c]	11	60	-.043 (.018)*	-.032 (.010)* *	$F(1, 58) = 10.58$	.002* *
TSST responder rate[c, d]	10	59	-.038 (.029)	-.001 (.001)	$F(1, 57) = 1.10$	.298
Recovery computation method:					$F(2, 56) = 0.14$	.874
C_{Min} (RC)	10	20	-.042 (.035)			
BLLV	10	20	-.028 (.036)	.014 (.040)		
Percentage Change	9	19	-.022 (.037)	.021 (.041)		

Note. #k = number of studies; #ES = number of effect sizes; Z_r = Fisher's Z correlation; *SE* = standard error; β_1 = estimated regression coefficient; *df* = degrees of freedom; AUC_I = area under the curve with respect to increase; AUC_G = area under the curve with respect to ground; CAR = cortisol awakening response; C_{Min} = minimum cortisol value; MaxMin = maximum value minus minimum value; Peak-Baseline = peak value minus baseline value; RC = reference category.
[a] For categorical moderators, a study can be represented in more than one moderator category if the effect sizes reported relate to different moderator categories. Thus, the sum of the number of studies (*k*) included in distinct categories belonging to the same moderator can exceed 12 (the total number of unique studies); [b] Omnibus test of all regression coefficients in the model; [c] continuous variable; [d] defined as percentage of participants showing a cortisol increase > 1.5 nmol/l from baseline to individual peak levels throughout the TSST.
* $p < .05$; * * $p < .01$; * ** $p < .001$

time points. Another recommendation, that is to monitor CAR sampling times objectively (Stalder et al., 2016), was also not commonly implemented (4 out of 11 studies). Together, these methodological issues may have lowered reliability of diurnal cortisol parameters. With regards to the TSST, on average, only half of the participants showed a significant cortisol increase, indicating successful stress induction. The observed relatively low responder rates question the extent to which stress reactivity has been reliably assessed in non-responders.

At the meta-analytic level, we focused on the most commonly employed measures of diurnal cortisol and the cortisol stress response (see Table 1), leaving aside the many more ways to compute these parameters (see e.g., Abelson et al., 2023; Khoury et al., 2015; Miller et al., 2018). In respect of recent recommendations, we included measures that were shown in a data-driven study to best reflect stress reactivity (MaxMin) and recovery (C_{Min}; Miller et al., 2018). Of note, our findings revealed that some indices of stress reactivity and recovery were interrelated, suggesting that disentangling both processes remains a challenge. While higher reactivity computed as AUC_I or Peak-Baseline was significantly related to *lower* (i.e., worse) recovery computed as BLLV, there were neither associations between the reactivity index MaxMin with recovery indices, nor between the recovery index Percentage Change with reactivity indices. These divergent patterns of associations question the validity of commonly employed measures of stress reactivity and recovery, and indicate the need for consensus on which ones to apply.

In addition, although the decision to include only healthy participants was made because HPA axis functioning is altered in the context of psychopathology (Buitelaar, 2013), broadening (meta-analytic) research on associations between diurnal and reactive HPA axis states to clinical populations will be needed to gain deeper understanding of HPA axis regulation in different states.

Related to our decision to focus on studies employing the TSST, the range of participants' mean age was restricted to school-aged children, adolescents, and young to middle-aged adults. The youngest participants' mean age among included studies was 7.2 years (Raffington et al., 2018), which corresponds to the fact that the modified, child-appropriate version of the TSST is employed from school age onwards (Seddon et al., 2020). The oldest participants' mean age among included studies was 45.4 year (Juster et al., 2011). While we did not find evidence for a moderating role of age in the association between diurnal cortisol and cortisol stress reactivity, we cannot generalize this finding to early childhood and late adulthood. Hence, it remains to be tested whether the association between diurnal and reactive cortisol changes as a function of developmental stage. This may in particular be the case in early childhood, during which brain regions involved in HPA axis regulation are undergoing maturation (Loman and Gunnar, 2010).

As a final note, from 59 eligible studies that had assessed both diurnal and reactive cortisol, only 15 authors were able to provide effect size data. A common reason reported by authors was having no access to the data anymore. Consequently, a selection bias might have occurred, which entails that the obtained meta-analytic results rely primarily on those studies of researchers who kept proper records of their data. As the meaningfulness of any meta-analysis increases with the number of studies included, we call for better long-term data storage and sharing among researchers.

4.2. Conclusions

Little work has been done to discover associations between diurnal and reactive HPA axis states. In sum, evidence that diurnal cortisol parameters would relate to acute cortisol stress reactivity and recovery is rather weak. Exceptions portray links between better stress recovery when computed as C_{Min} with lower total diurnal cortisol output and steeper diurnal slopes, which are considered indicators of "adaptive" HPA axis regulation (Adam et al., 2017). Still, it remains unclear whether these reflect "true" associations or whether they arise due to conflating of the stress recovery measure C_{Min} with diurnal measures. Our findings further support the notion that an AUC_G stress measure is not representing stress reactivity accurately, and advise caution in the use of terminology. Further, the lack of an association among commonly employed stress recovery parameters indicates that they might measure different aspects of the stress response. More research is needed to identify under which circumstances diurnal cortisol parameters relate to stress reactivity and recovery, which may depend on the computation method used to derive diurnal and stress response parameters, the stress task, and sample characteristics such as psychopathology.

Other information

This meta-analysis was pre-registered on the 28th of November 2022 through a web-based protocol on the International Prospective Register of Systematic Reviews (PROSPERO; Booth et al., 2012) before study selection (registration number: CRD42022372181). Meta-analytic data sets and corresponding scripts for statistical analyses are available on OSF (https://osf.io/t4dxa/).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Wesarg-Menzel Christiane: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft. **Engert Veronika:** Supervision, Writing – review & editing. **Marheinecke Ruth:** Data curation, Validation, Writing – review & editing. **Staaks Janneke:** Methodology, Writing – review & editing.

Declaration of Competing Interest

None.

Acknowledgements

We would like to thank all authors who responded to our data requests: Magdalena Degering, Mark A. Ellenbogen, Elissa Epel, Brittany Evans, Brigitte M. Kudielka, Sonia J. Lupien, Rachel G. Lucas-Thompson, Robert-Paul Juster, Laurel Raffington, Yee Lee Shing, A. Janet Tomiyama, Petra H. Wirtz, and Yujiro Yamanaka.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.psyneuen.2024.106976.

References

References with an asterisk ('*') were included in the meta-analysis.

Abelson, J.L., Sanchez, B.N., Mayer, S.E., Briggs, H., Liberzon, I., Rajaram, N., 2023. Do diurnal salivary cortisol curves carry meaningful information about the regulatory biology of the HPA axis in healthy humans? Psychoneuroendocrinology 150, 106031. https://doi.org/10.1016/j.psyneuen.2023.106031.

Adam, E.K., Quinn, M.E., Tavernier, R., McQuillan, M.T., Dahlke, K.A., Gilbert, K.E., 2017. Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis. Psychoneuroendocrinology 83, 25–41. https://doi.org/10.1016/j.psyneuen.2017.05.018.

Assink, M., Wibbelink, C., 2016. Fitting three-level meta-analytic models in R: a step-by-step tutorial. Quant. Methods Psychol. 12 https://doi.org/10.20982/tqmp.12.3.p154.

Buitelaar, J.K., 2013. The role of the HPA-axis in understanding psychopathology: cause, consequence, mediator, or moderator? Eur. Child Adolesc. Psychiatry 22 (7), 387–389. https://doi.org/10.1007/s00787-013-0441-7.

Buske-Kirschbaum, A., Jobst, S., Wustmans, A., Kirschbaum, C., Rauh, W., Hellhammer, D., 1997. Attenuated free cortisol response to psychosocial stress in

children with atopic dermatitis. Psychosom. Med 59 (4), 419–426. https://doi.org/10.1097/00006842-199707000-00012.
Charmandari, E., Tsigos, C., Chrousos, G., 2005. Endocrinology of the stress response. Annu Rev. Physiol. 67, 259–284. https://doi.org/10.1146/annurev.physiol.67.040403.120816.
Cheung, M.W.-L., 2015. metaSEM: an R package for meta-analysis using structural equation modeling [Methods]. Front. Psychol. 5 https://doi.org/10.3389/fpsyg.2014.01521.
Crişan, L.G., Vulturar, R., Miclea, M., Miu, A.C., 2016. Reactivity to social stress in subclinical social anxiety: emotional experience, cognitive appraisals, behavior, and physiology. Front Psychiatry 7, 5. https://doi.org/10.3389/fpsyt.2016.00005.
**Degering, M., Linz, R., Puhlmann, L.M.C., Singer, T., Engert, V., 2023. Revisiting the stress recovery hypothesis: Differential associations of cortisol stress reactivity and recovery after acute psychosocial stress with markers of long-term stress and health. Brain Behav. Immun. Health 28, 100598. https://doi.org/10.1016/j.bbih.2023.100598.
De Kloet, E.R., Sutanto, W., Rots, N., van Haarst, A., van den Berg, D., Oitzl, M., van Eekelen, A., Voorhuis, D., 1991. Plasticity and function of brain corticosteroid receptors during aging. Acta Endocrinol. 125 (*Suppl 1*), 65–72 http://europepmc.org/abstract/MED/1801504.
DeMorrow, S., 2018. Role of the hypothalamic-pituitary-adrenal axis in health and disease. Int J. Mol. Sci. 19 (4) https://doi.org/10.3390/ijms19040986.
Dickerson, S.S., Kemeny, M.E., 2004. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol. Bull. 130 (3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355.
**Ellenbogen, M.A., Hodgins, S., Walker, C.D., Couture, S., Adam, S., 2006. Daytime cortisol and stress reactivity in the offspring of parents with bipolar disorder. Psychoneuroendocrinology 31 (10), 1164–1180. https://doi.org/10.1016/j.psyneuen.2006.08.004.
**Evans, B.E., Greaves-Lord, K., Euser, A.S., Franken, I.H., Huizink, A.C., 2013. Cortisol levels in children of parents with a substance use disorder. Psychoneuroendocrinology 38 (10), 2109–2120. https://doi.org/10.1016/j.psyneuen.2013.03.021.
Fairchild, G., van Goozen, S.H., Stollery, S.J., Brown, J., Gardiner, J., Herbert, J., Goodyer, I.M., 2008. Cortisol diurnal rhythm and stress reactivity in male adolescents with early-onset or adolescence-onset conduct disorder. Biol. Psychiatry 64 (7), 599–606. https://doi.org/10.1016/j.biopsych.2008.05.022.
Fekedulegn, D.B., Andrew, M.E., Burchfiel, C.M., Violanti, J.M., Hartley, T.A., Charles, L. E., Miller, D.B., 2007. Area under the curve and other summary indicators of repeated waking cortisol measurements. Psychosom. Med 69 (7), 651–659. https://doi.org/10.1097/PSY.0b013e31814c405c.
Fries, E., Dettenborn, L., Kirschbaum, C., 2009. The cortisol awakening response (CAR): facts and future directions. Int J. Psychophysiol. 72 (1), 67–73. https://doi.org/10.1016/j.ijpsycho.2008.03.014.
Gaab, J., Rohleder, N., Nater, U.M., Ehlert, U., 2005. Psychological determinants of the cortisol stress response: the role of anticipatory cognitive appraisal. Psychoneuroendocrinology 30 (6), 599–610. https://doi.org/10.1016/j.psyneuen.2005.02.001.
Gerber, M., Imboden, C., Beck, J., Brand, S., Colledge, F., Eckert, A., Holsboer-Trachsler, E., Puhse, U., Hatzinger, M., 2020. Effects of aerobic exercise on cortisol stress reactivity in response to the trier social stress test in inpatients with major depressive disorders: a randomized controlled trial. J. Clin. Med 9 (5). https://doi.org/10.3390/jcm9051419.
Gunn, P.J., Middleton, B., Davies, S.K., Revell, V.L., Skene, D.J., 2016. Sex differences in the circadian profiles of melatonin and cortisol in plasma and urine matrices under constant routine conditions. Chrono-.-. Int 33 (1), 39–50. https://doi.org/10.3109/07420528.2015.1112396.
Halbeisen, G., Domes, G., Walther, E., 2023. Is stress colorblind? Exploring endocrine stress responses in intergroup contexts using a virtual reality-based Trier Social Stress Test (TSST-VR). Psychoneuroendocrinology 147, 105970. https://doi.org/10.1016/j.psyneuen.2022.105970.
Hellhammer, J., Fries, E., Schweisthal, O.W., Schlotz, W., Stone, A.A., Hagemann, D., 2007. Several daily measurements are necessary to reliably assess the cortisol rise after awakening: state- and trait components. Psychoneuroendocrinology 32 (1), 80–86. https://doi.org/10.1016/j.psyneuen.2006.10.005.
Hostinar, C.E., Johnson, A.E., Gunnar, M.R., 2015. Parent support is less effective in buffering cortisol stress reactivity for adolescents compared to children. Dev. Sci. 18 (2), 281–297. https://doi.org/10.1111/desc.12195.
Hulett, J.M., Fessele, K.L., Clayton, M.F., Eaton, L.H., 2019. Rigor and reproducibility: a systematic review of salivary cortisol sampling and reporting parameters used in cancer survivorship research. Biol. Res Nurs. 21 (3), 318–334. https://doi.org/10.1177/1099800419835321.
Jacobson, L., Sapolsky, R., 1991. The role of the hippocampus in feedback regulation of the hypothalamic-pituitary-adrenocortical axis. Endocr. Rev. 12 (2), 118–134. https://doi.org/10.1210/edrv-12-2-118.
Juster, R.P., Perna, A., Marin, M.F., Sindi, S., Lupien, S.J., 2012. Timing is everything: anticipatory stress dynamics among cortisol and blood pressure reactivity and recovery in healthy adults. Stress 15 (6), 569–577. https://doi.org/10.3109/10253890.2012.661494.
**Juster, R.P., Raymond, C., Desrochers, A.B., Bourdon, O., Durand, N., Wan, N., Pruessner, J.C., Lupien, S.J., 2016. Sex hormones adjust "sex-specific" reactive and diurnal cortisol profiles. Psychoneuroendocrinology 63, 282–290. https://doi.org/10.1016/j.psyneuen.2015.10.012.
**Juster, R.P., Sindi, S., Marin, M.F., Perna, A., Hashemi, A., Pruessner, J.C., Lupien, S.J., 2011. A clinical allostatic load index is associated with burnout symptoms and hypocortisolemic profiles in healthy workers. Psychoneuroendocrinology 36 (6), 797–805. https://doi.org/10.1016/j.psyneuen.2010.11.001.
**Kerr, P., Lupien, S., Juster, R.P., 2020. Rx risk or resistance? Psychotropic medication use in relation to physiological and psychosocial functioning of psychiatric hospital workers. Psychoneuroendocrinology 115, 104634. https://doi.org/10.1016/j.psyneuen.2020.104634.
Khoury, J.E., Gonzalez, A., Levitan, R.D., Pruessner, J.C., Chopra, K., Basile, V.S., Masellis, M., Goodwill, A., Atkinson, L., 2015. Summary cortisol reactivity indicators: Interrelations and meaning. Neurobiol. Stress 2, 34–43. https://doi.org/10.1016/j.ynstr.2015.04.002.
Kidd, T., Carvalho, L.A., Steptoe, A., 2014. The relationship between cortisol responses to laboratory stress and cortisol profiles in daily life. Biol. Psychol. 99 (100), 34–40. https://doi.org/10.1016/j.biopsycho.2014.02.010.
Kirschbaum, C., Hellhammer, D.H., 1989. Salivary cortisol in psychobiological research: an overview. Neuropsychobiology 22 (3), 150–169. https://doi.org/10.1159/000118611.
Kirschbaum, C., Pirke, K.M., Hellhammer, D.H., 1993. The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology 28 (1-2), 76–81. https://doi.org/10.1159/000119004.
Kirschbaum, C., Klauer, T., Filipp, S.H., Hellhammer, D.H., 1995. Sex-specific effects of social support on cortisol and subjective responses to acute psychological stress. Psychosom. Med 57 (1), 23–31. https://doi.org/10.1097/00006842-199501000-00004.
Kudielka, B.M., Schommer, N.C., Hellhammer, D.H., Kirschbaum, C., 2004. Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. Psychoneuroendocrinology 29 (8), 983–992. https://doi.org/10.1016/j.psyneuen.2003.08.009.
Law, R., Hucklebridge, F., Thorn, L., Evans, P., Clow, A., 2013. State variation in the cortisol awakening response. Stress 16 (5), 483–492. https://doi.org/10.3109/10253890.2013.817552.
Lipsey, M.W., Wilson, D.B., 2001. Practical meta-analysis. Sage Publications, Inc.
Liu, J.J.W., Ein, N., Peck, K., Huang, V., Pruessner, J.C., Vickers, K., 2017. Sex differences in salivary cortisol reactivity to the Trier Social Stress Test (TSST): a meta-analysis. Psychoneuroendocrinology 82, 26–37. https://doi.org/10.1016/j.psyneuen.2017.04.007.
Loman, M.M., Gunnar, M.R., 2010. Early experience and the development of stress reactivity and regulation in children. Neurosci. Biobehav. Rev. 34 (6), 867–876.
Lopez-Duran, N.L., Mayer, S.E., Abelson, J.L., 2014. Modeling neuroendocrine stress reactivity in salivary cortisol: adjusting for peak latency variability. Stress 17 (4), 285–295. https://doi.org/10.3109/10253890.2014.915517.
**Lucas-Thompson, R.G., Henry, K.L., McKernan, C.J., 2018. Is cortisol production in response to an acute stressor associated with diurnal cortisol production during adolescence? Dev. Psychobiol. 60 (4), 449–457. https://doi.org/10.1002/dev.21593.
**Lupien, S.J., Leclaire, S., Majeur, D., Raymond, C., Jean Baptiste, F., Giguere, C.E., 2022. Doctor, I am so stressed out!' A descriptive study of biological, psychological, and socioemotional markers of stress in individuals who self-identify as being 'very stressed out' or 'zen. '. *Neurobiol. Stress* 18, 100454. https://doi.org/10.1016/j.ynstr.2022.100454.
**Malanchini, M., Engelhardt, L.E., Raffington, L.A., Sabhlok, A., Grotzinger, A.D., Briley, D.A., Madole, J.W., Freis, S.M., Patterson, M.W., Harden, K.P., Tucker-Drob, E.M., 2021. Weak and uneven associations of home, neighborhood, and school environments with stress hormone output across multiple timescales. Mol. Psychiatry 26 (9), 4823–4838. https://doi.org/10.1038/s41380-020-0747-z.
Mikneviciute, G., Pulopulos, M.M., Allaert, J., Armellini, A., Rimmele, U., Kliegel, M., Ballhausen, N., 2023. Adult age differences in the psychophysiological response to acute stress. Psychoneuroendocrinology 153, 106111. https://doi.org/10.1016/j.psyneuen.2023.106111.
Miller, R., Wojtyniak, J.G., Weckesser, L.J., Alexander, N.C., Engert, V., Lehr, T., 2018. How to disentangle psychobiological stress reactivity and recovery: A comparison of model-based and non-compartmental analyses of cortisol concentrations. Psychoneuroendocrinology 90, 194–210. https://doi.org/10.1016/j.psyneuen.2017.12.019.
Nierop, A., Bratsikas, A., Klinkenberg, A., Nater, U.M., Zimmermann, R., Ehlert, U., 2006. Prolonged salivary cortisol recovery in second-trimester pregnant women and attenuated salivary alpha-amylase responses to psychosocial stress in human pregnancy. J. Clin. Endocrinol. Metab. 91 (4), 1329–1335. https://doi.org/10.1210/jc.2005-1816.
Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A., 2016. Rayyan—a web and mobile app for systematic reviews. Syst. Rev. 5 (1), 210 https://doi.org/10.1186/s13643-016-0384-4.
Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hrobjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372, n71. https://doi.org/10.1136/bmj.n71.
Papadimitriou, A., Priftis, K.N., 2009. Regulation of the hypothalamic-pituitary-adrenal axis. Neuroimmunomodulation 16 (5), 265–271. https://doi.org/10.1159/000216184.
Pruessner, J.C., Wolf, O.T., Hellhammer, D.H., Buske-Kirschbaum, A., von Auer, K., Jobst, S., Kaspers, F., Kirschbaum, C., 1997. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. Life Sci. 61 (26), 2539–2549. https://doi.org/10.1016/s0024-3205(97)01008-4.
Pruessner, J.C., Kirschbaum, C., Meinlschmid, G., Hellhammer, D.H., 2003. Two formulas for computation of the area under the curve represent measures of total

hormone concentration versus time-dependent change. Psychoneuroendocrinology 28 (7), 916–931. https://doi.org/10.1016/s0306-4530(02)00108-7.

**Raffington, L., Malanchini, M., Grotzinger, A.D., Madole, J.W., Engelhardt, L.E., Sabhlok, A., Youn, C., Patterson, M.W., Harden, K.P., Tucker-Drob, E.M., 2022. An in-laboratory stressor reveals unique genetic variation in child cortisol output. Dev. Psychol. 58 (10), 1832–1848. https://doi.org/10.1037/dev0001393.

**Raffington, L., Prindle, J., Keresztes, A., Binder, J., Heim, C., Shing, Y.L., 2018. Blunted cortisol stress reactivity in low-income children relates to lower memory function. Psychoneuroendocrinology 90, 110–121. https://doi.org/10.1016/j.psyneuen.2018.02.002.

Ross, K.M., Murphy, M.L.M., Adam, E.K., Chen, E., Miller, G.E., 2014. How stable are diurnal cortisol activity indices in healthy individuals? Evidence from three multi-wave studies. Psychoneuroendocrinology 39, 184–193. https://doi.org/10.1016/j.psyneuen.2013.09.016.

Sandner, M., Lois, G., Streit, F., Zeier, P., Kirsch, P., Wust, S., Wessa, M., 2020. Investigating individual stress reactivity: High hair cortisol predicts lower acute stress responses. Psychoneuroendocrinology 118, 104660. https://doi.org/10.1016/j.psyneuen.2020.104660.

Schlotz, W., Hammerfald, K., Ehlert, U., Gaab, J., 2011. Individual differences in the cortisol response to stress in young healthy men: testing the roles of perceived stress reactivity and threat appraisal using multiphase latent growth curve modeling. Biol. Psychol. 87 (2), 257–264. https://doi.org/10.1016/j.biopsycho.2011.03.005.

Schmidt-Reinwald, A., Pruessner, J.C., Hellhammer, D.H., Federenko, I., Rohleder, N., Schurmeyer, T.H., Kirschbaum, C., 1999. The cortisol response to awakening in relation to different challenge tests and a 12-hour cortisol rhythm. Life Sci. 64 (18), 1653–1660. https://doi.org/10.1016/s0024-3205(99)00103-4.

Schoorl, J., Rijn, S.V., Wied, M., van Goozen, S., Swaab, H., 2016. The role of anxiety in cortisol stress response and cortisol recovery in boys with oppositional defiant disorder/conduct disorder. Psychoneuroendocrinology 73, 217–223. https://doi.org/10.1016/j.psyneuen.2016.08.007.

Seddon, J.A., Rodriguez, V.J., Provencher, Y., Raftery-Helmer, J., Hersh, J., Labelle, P.R., Thomassin, K., 2020. Meta-analysis of the effectiveness of the Trier Social Stress Test in eliciting physiological stress responses in children and adolescents. Psychoneuroendocrinology 116, 104582.

Segerstrom, S.C., Boggero, I.A., Smith, G.T., Sephton, S.E., 2014. Variability and reliability of diurnal cortisol in younger and older adults: implications for design decisions. Psychoneuroendocrinology 49, 299–309. https://doi.org/10.1016/j.psyneuen.2014.07.022.

Shenk, C.E., Felt, J.M., Ram, N., O'Donnell, K.J., Sliwinski, M.J., Pokhvisneva, I., Benson, L., Meaney, M.J., Putnam, F.W., Noll, J.G., 2022. Cortisol trajectories measured prospectively across thirty years of female development following exposure to childhood sexual abuse: Moderation by epigenetic age acceleration at midlife. Psychoneuroendocrinology 136, 105606. https://doi.org/10.1016/j.psyneuen.2021.105606.

Simons, S.S., Cillessen, A.H., de Weerth, C., 2017. Associations between circadian and stress response cortisol in children. Stress 20 (1), 52–58. https://doi.org/10.1080/10253890.2016.1276165.

Spencer, R.L., Deak, T., 2017. A users guide to HPA axis research. Physiol. Behav. 178, 43–65. https://doi.org/10.1016/j.physbeh.2016.11.014.

Staaks, J., 2020. 1. Zotero (recommended) deduplication, version 4. Syst Rev. Search Support. ⟨https://osf.io/yh3x⟩e/).

Stalder, T., Kirschbaum, C., Kudielka, B.M., Adam, E.K., Pruessner, J.C., Wust, S., Dockray, S., Smyth, N., Evans, P., Hellhammer, D.H., Miller, R., Wetherell, M.A., Lupien, S.J., Clow, A., 2016. Assessment of the cortisol awakening response: Expert consensus guidelines. Psychoneuroendocrinology 63, 414–432. https://doi.org/10.1016/j.psyneuen.2015.10.010.

Tsigos, C., Chrousos, G.P., 2002. Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress. J. Psychosom. Res 53 (4), 865–871. https://doi.org/10.1016/s0022-3999(02)00429-4.

Van den Noortgate, W., López-López, J.A., Marín-Martínez, F., Sánchez-Meca, J., 2013. Three-level meta-analysis of dependent effect sizes. Behav. Res. Methods 45 (2), 576–594. https://doi.org/10.3758/s13428-012-0261-6.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36 (3), 1–48. https://doi.org/10.18637/jss.v036.i03.

de Weerth, C., Zijlmans, M.A., Mack, S., Beijers, R., 2013. Cortisol reactions to a social evaluative paradigm in 5- and 6-year-old children. Stress 16 (1), 65–72. https://doi.org/10.3109/10253890.2012.684112.

**Wirtz, P.H., Ehlert, U., Emini, L., Suter, T., 2008. Higher body mass index (BMI) is associated with reduced glucocorticoid inhibition of inflammatory cytokine production following acute psychosocial stress in men. Psychoneuroendocrinology 33 (8), 1102–1110. https://doi.org/10.1016/j.psyneuen.2008.05.002.

**Wirtz, P.H., Elsenbruch, S., Emini, L., Rudisuli, K., Groessbauer, S., Ehlert, U., 2007. Perfectionism and the cortisol response to psychosocial stress in men. Psychosom. Med 69 (3), 249–255. https://doi.org/10.1097/PSY.0b013e318042589e.

**Wolfram, M., Bellingrath, S., Feuerhahn, N., Kudielka, B.M., 2013. Cortisol responses to naturalistic and laboratory stress in student teachers: comparison with a non-stress control day. Stress Health 29 (2), 143–149. https://doi.org/10.1002/smi.2439.

**Yamanaka, Y., Motoshima, H., Uchida, K., 2019. Hypothalamic-pituitary-adrenal axis differentially responses to morning and evening psychological stress in healthy subjects. Neuropsychopharmacol. Rep. 39 (1), 41–47. https://doi.org/10.1002/npr2.12042.

Zänkert, S., Bellingrath, S., Wust, S., Kudielka, B.M., 2019. HPA axis responses to psychological challenge linking stress and disease: what do we know on sources of intra- and interindividual variability? Psychoneuroendocrinology 105, 86–97. https://doi.org/10.1016/j.psyneuen.2018.10.027.

Zorn, J.V., Schur, R.R., Boks, M.P., Kahn, R.S., Joels, M., Vinkers, C.H., 2017. Cortisol stress reactivity across psychiatric disorders: a systematic review and meta-analysis. Psychoneuroendocrinology 77, 25–36. https://doi.org/10.1016/j.psyneuen.2016.11.036.