Frontiers in Neuroendocrinology 68 (2023) 101050

Contents lists available at ScienceDirect

Frontiers in Neuroendocrinology

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

Multi-systemic evaluation of biological and emotional responses to the Trier Social Stress Test: A meta-analysis and systematic review

Idy S.C. Man [a,b,#], Robin Shao [a,b,k,#], W.K. Hou [c,d], Shirley Xin Li [a,e], Fiona Yan Liu [f], Maggy Lee [g], Yun Kwok Wing [h], Suk-yu Yau [i,j,*], Tatia M.C. Lee [a,b,*]

[a] State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, China
[b] Laboratory of Neuropsychology and Human Neuroscience, The University of Hong Kong, Hong Kong, China
[c] Department of Psychology, The Education University of Hong Kong, Hong Kong, China
[d] Centre for Psychosocial Health, The Education University of Hong Kong, Hong Kong, China
[e] Sleep Research Clinic and Laboratory, Department of Psychology, The University of Hong Kong, Hong Kong, China
[f] Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China
[g] Department of Sociology, The University of Hong Kong, Hong Kong, China
[h] Li Chiu Kong Family Sleep Assessment Unit, Department of Psychiatry, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China
[i] Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China
[j] Mental Health Research Center, The Hong Kong Polytechnic University, Hong Kong, China
[k] Department of Affective Disorders, The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou, China

ARTICLE INFO

Keywords:
Meta-analysis
Acute Stress
TSST
Biomarker
Cortisol
Emotion

ABSTRACT

Humans experience multiple biological and emotional changes under acute stress. Adopting a multi-systemic approach, we summarized 61 studies on healthy people's endocrinological, physiological, immunological and emotional responses to the Trier Social Stress Test. We found salivary cortisol and negative mood states were the most sensitive markers to acute stress and recovery. Biomarkers such as heart rate and salivary alpha-amylase also showed sensitivity to acute stress, but the numbers of studies were small. Other endocrinological (e.g., dehydroepiandrosterone), inflammatory (C-Reactive Protein, Interleukin-6) and physiological (e.g., skin conductance level) measures received modest support as acute stress markers. Salivary cortisol showed some associations with mood measures (e.g., state anxiety) during acute stress and recovery, and heart rate showed preliminary positive relationship with calmness ratings during response to TSST, but the overall evidence was mixed. While further research is needed, these findings provide updated and comprehensive knowledge on the integrated psychobiological response profiles to TSST.

1. Introduction

It is widely acknowledged that stress responses entail multi-faceted biological and psychological processes which interact in intricate ways (e.g., McEwen, 1998; Schneiderman et al., 2005; Juster et al., 2010). On one hand, multiple neuroendocrinological, immunological and cardiovascular systems interact among each other and collectively determine the stress allostatic load (Juster et al., 2010). Therefore, stress outcomes are results of the confluence of changes accumulated across multiple biological systems (Seeman et al., 2001). On the other hand, biology and emotion are inseparable realms in stress responses. According to the classic James-Lange theory, emotions emerge from physiological changes and are closely related to visceral functions (Fehr and Stern, 1970). Overlapping brain centers such as the prefrontal-limbic circuitries orchestrate both biological and emotional responses to stress, providing the neural bases for their tight interactions (Kern et al., 2008; Fanselow and Dong, 2010; Flandreau et al., 2012; Lee et al., 2012). Thus, to understand the global profile of acute stress reactivity and recovery, it is crucial to adopt a multi-systemic approach, through simultaneously and comprehensively examine both biological and emotional responses, *as well as* the biology-emotion interrelations.

Stress responses are primarily mediated via activating the

* Corresponding authors at: Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China (S.-Y. Yau). State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, China (Tatia M.C. Lee).
E-mail addresses: sonata.yau@polyu.edu.hk (S.-y. Yau), tmclee@hku.hk (T.M.C. Lee).
# These authors contributed equally to the manuscript.

https://doi.org/10.1016/j.yfrne.2022.101050
Received 31 May 2022; Received in revised form 28 September 2022; Accepted 13 November 2022
Available online 18 November 2022

hypothalamic–pituitaryadrenal (HPA) axis and the sympathetic-adrenomedullar (SAM) system (See (Schneiderman et al., 2005) for detailed review). On one hand, the hypothalamus in the HPA axis produces corticotropin releasing factors, which induce the release of adrenocorticotropic hormone (ACTH), and further stimulate the production and secretion of cortisol. Hormonal responses are relatively slower and can last longer. For example, cortisol responses are typically delayed relative to stress onset by around 15 min. Cortisol reaction corresponding to the time immediately before or shortly after stressor onset is generally considered as anticipatory stress, whereas cortisol response corresponding to later parts of the stressor is considered to reflect reactive stress (Mikolajczak et al., 2008). On the other hand, the SAM axis secretes adrenaline and noradrenaline to provide rapid physiological adaptation to stress, including the elevation of heart rate (HR), breathing rate and blood pressure (BP). Given the swift changes of these markers, the cardiovascular responses are often monitored during the stress process to capture the immediate stress reaction profile. Other well-documented physiological and hormonal responses to acute stress include the secretion of salivary alpha-amylase (sAA), an enzyme that facilitates digestion in the oral cavity (Petrakova et al., 2015); elevated level of dehydroepiandrosterone (DHEA), an androgen precursor secreted by the adrenal gland (Lennartsson et al., 2012); and changes in BP, skin conductance level (SCL) and heart-rate variability (HRV) (Wemm and Wulfert, 2017; Berntson and Cacioppo, 2004). In addition, the immune systems are also sensitive to stress. Exposure to stressors have been linked with prolonged proinflammatory cytokine releases and increased levels of circulating cytokines, including interleukin-6 (IL-6) and C-reactive protein (CRP) (see (Steptoe et al., 2007) for systematic review).

Besides the biological changes to acute stress, mood states are often self-reported pre- and post-stressor to record the emotional changes. Individuals exposed to stress generally report a subjectively negative experience with increased perceived stress and anxiety [e.g., Allen et al., 2014; Merz and Wolf, 2015]. However, to our best knowledge, there has been no systematic review or *meta*-analysis on the mood state changes in response to acute stress.

Although a number of previous reviews examined biological responses to acute stress, only one narrative review examined multiple biological and psychological markers (Allen et al., 2014). Also, most existing *meta*-analyses on biomarkers of acute stress examined only one specific type of marker, such as cortisol (Helminen et al., 2019; Liu et al., 2017 Aug), DHEA (Dutheil et al., 2021), salivary inflammatory markers (Steptoe et al., 2007; Szabo et al., 2020), and blood pressure (Gasperin et al., 2009). Moreover, few reviews have examined the recovery patterns of acute stress responses, or the association of biological and psychological response changes during acute stress response and recovery (Dickerson and Kemeny, 2004; Dickerson, 2008).

Therefore, in this systematic review and *meta*-analysis, we aimed to comprehensively summarize existing empirical evidence on biological, endocrinological/hormonal and emotional responses to the most commonly adopted acute stress paradigm, the Trier Social Stress Test (TSST) (Kirschbaum et al., 1993), encompassing both stress reactivity and recovery. Importantly, we also reviewed the interrelations between changes in all common biomarkers and both positive and negative emotions across the TSST, in order to provide an updated account on the biology-emotion association in acute stress outcome.

2. Methods

This review was conducted following the Preferred Reporting Items for Systematic Reviews and meta-Analyses (PRISMA) guidelines (Moher et al., 2009). The review protocol was pre-registered at PROSPERO, International prospective register of systematic reviews (Reference number: CRD42021228639). The PRISMA checklist can be found in Supplementary Material 1.

2.1. Search strategy and selection criteria

Searches of the current study were conducted using PsycINFO, PubMed, Web of Science, and PsycArticles electronic databases. The last search was completed on March 19, 2021, with the following keywords: (Physiological OR biological OR cortisol) AND (affective OR emotional OR "self-report" OR health OR well*) AND "Trier social stress test". Only studies that were fully written in English were considered.

This *meta*-analysis included adult healthy samples only. Any subjects with records of psychiatric, neurological or endocrinological disorders were excluded. Due to the confounding effects of fluctuations in hormonal status upon stress responses, pregnant and postmenopausal women were excluded. However, due to the high interest on the effect of stress on postmenopausal women, we conducted supplementary analyses on that population (Supplementary Material 5). Moreover, to minimize the confounding effect of aging on stress response, old-age populations (>65 years old) were excluded.

Inclusion and exclusion criteria are detailed in Supplementary Material 2 Table S1A.

2.2. Data selection, extraction and coding

2.2.1. Data selection

The literature search identified 2581 articles after initial search. After duplicates were removed, a total of 1465 unique studies were identified. Study selection and screening procedure were informed by PRISMA guidelines (see Fig. 1 for the PRISMA flow chart). Two independent reviewers screened the titles and abstracts during the first-stage screening. Studies that met any of the exclusion criteria were screened out. The remaining studies were passed to the second stage for assessment of full text eligibility to resolve any remaining ambiguity. First-stage study screening yielded high level of inter-rater agreement (inter-rater agreement = 97.47 %, Cohen's kappa = 0.90). Any discrepancies in the screening results were resolved by discussion, and unresolved discrepancies were referred to a third reviewer.

During Stage 1, out of the 1465 records, 1254 records were removed: 139 papers were removed because they were reviews rather than original research papers; 5 were excluded because English full text was not available; 212 papers were excluded due to their sample ages not in the adulthood range (i.e., children or adolescents); 58 were excluded due to their samples not healthy (e.g., cancer patients, current or remitted psychological disorders); 26 were excluded due to their female samples being in pregnancy, post-partum or post-menopausal stage; 270 were excluded due to the use of not standard TSST paradigm (e.g., multi-subject setting, absence of panel audience, virtual reality setting, no speech or mental arithmetic task; given our focus was not to evaluate the impact of different versions of TSST on responses, we only included the standard TSST tasks); 199 studies were excluded since they administered concurrent or prior pharmacological, cognitive, emotional or mindfulness-based interventions; 24 were excluded due to use of concurrent stressors such as pain, fear-conditioned stimuli and other cognitive tasks; 13 were excluded due to lack of biomarkers other than neuroimaging measures; 308 were excluded due to lack of subjective emotional state measures.

A total of 211 studies were assessed in full text in second stage. Full-text eligibility assessment yielded high level of interrater agreement (inter-rater agreement = 96.21 %, Cohen's kappa = 0.92). All discrepancies in the eligibility assessment were resolved through discussion. Finally, 61 papers published between 2005 and 2020 were channeled to data extraction (see Fig. 1 for details about stage 2 exclusion). Of these, 58 papers contained variables with sufficient number of datasets for conducting *meta*-analyses . The remaining 3 papers only contained variables that could be narratively reviewed. Please note that the total number of papers included in *meta*-analysis and in narrative review added up to >61 since certain papers contained both variable(s) eligible for *meta*-analysis, and variable(s) only eligible for narrative review.

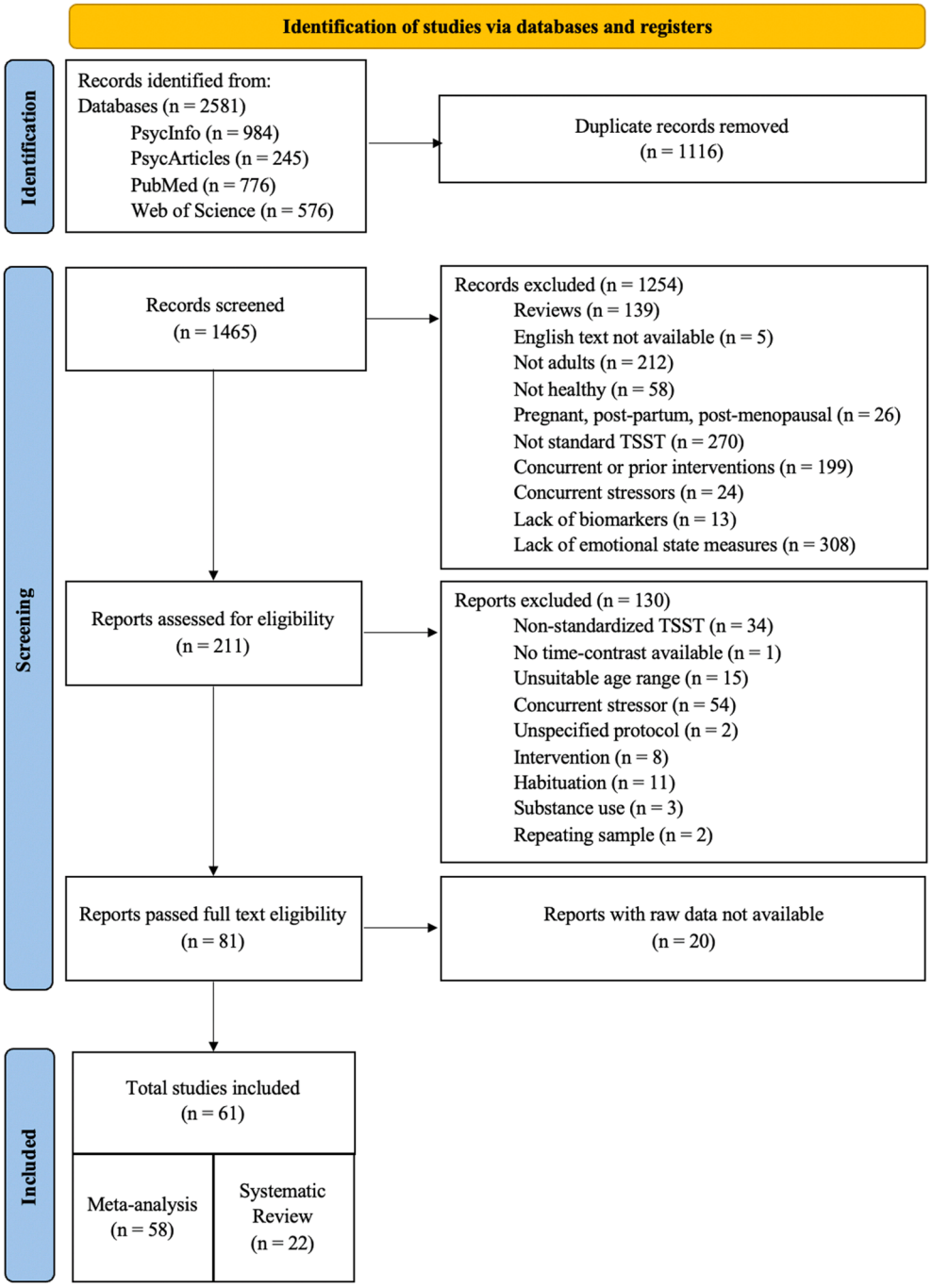

Fig. 1. The PRISMA flowchart for literature searching and screening. In total, 61 papers published between 2005 and 2020 were included. Of these, 58 papers contained variables with sufficient number of datasets for conducting *meta*-analyses. The remaining 3 papers only contained variables that could be narratively reviewed. Note that the number of papers included in *meta*-analysis and in narrative review added up to greater than 61 since certain papers contained both variable (s) eligible for *meta*-analysis, and variable(s) only eligible for narrative review.

2.2.2. *Data extraction*

Data extraction was performed by two authors (I.S.C.M. and R.S.). The mean and standard deviations of each measure were extracted for effect size calculation. If relevant studies did not provide sufficient data, the authors were contacted through e-mail to request for data sharing. For data not reported and not provided by the authors but figures were presented, a graphical extraction tool (Getdata Graph Digitizer version 2.26) was used to extract datapoints as numerical values (https://getdata-graph-digitizer.com). meta-analysis was performed for each measure for a given time contrast when ≥ 5 studies were available, while systematic review was conducted to summarize the findings for studies with $N_{study} < 5$.

Meta-analyses were run on seven measures in total. Five of them were biomarkers, with four endocrinological measures: salivary cortisol (SC), sAA, plasma cortisol (PC), and plasma adrenocorticotropic hormone (ACTH) and one physiological measure HR. The other two were positive (PA) and negative (NA) emotion state measures. Please refer to Supplement 2 Table 1B for details about emotion state measurements.

Eight variables were reviewed systematically, including 3 salivary biomarkers (DHEA, IL-6, CRP), 2 blood biomarkers (IL-6, prolactin) and 3 other physiological markers (SCL, BP and HRV).

2.2.3. *Data coding*

Since the data collection timepoints vary largely between markers and among studies, a standardized protocol was adopted to label the timeline of the experimental procedures. All time contrasts and labels are listed in detail in Supplementary tables (Supplementary Material 2 Tables S2A-2E, Tables S3A-3B). Briefly, time points were characterized

Table 1
Meta-analysis of biomarkers in response to TSST.

Biomarkers time contrasts	n	SMD (95 % CI)	*p* value	Heterogeneity			Egger's test *t* stat (*p* value)
				Q stat (df; *p*)	τ^2	I^2	
SC							
T_{pre} vs T_{post}	1050	0.82 (0.69 to 0.95)	<0.0001	201.04 (28; $p < 0.0001$)	0.058	86.1 %	−1.13 ($p = 0.270$)
T_{pre} vs T_{peak}	1023	0.85 (0.73 to 0.97)	<0.0001	242.69 (27; $p < 0.0001$)	0.037	88.9 %	−2.00 ($p = 0.056$)
T_{post} vs T_{rec1}	852	−0.31 (−0.50 to −0.13)	0.002	27.48 (20; $p = 0.122$)	0.098	27.2 %	−0.71 ($p = 0.485$)
T_{post} vs T_{rec2}	362	−0.70 (−0.89 to −0.50)	<0.0001	8.07 (8; $p = 0.427$)	0.049	0.8 %	N/A
T_{peak} vs T_{rec1}	712	−0.98 (−1.04 to −0.91)	<0.0001	21.56 (19; $p = 0.307$)	0.006	11.9 %	−0.50 ($p = 0.626$)
T_{peak} vs T_{rec2}	303	−0.57 (−0.72 to −0.41)	<0.0001	0.99 (7; $p = 0.995$)	0	0.00 %	N/A
sAA							
T_{pre} vs T_{post}	314	1.18 (1.10 to 1.26)	<0.0001	0.02 (8; $p = 1.000$)	0	0.00 %	N/A
T_{post} vs T_{rec1}	199	−0.36 (−0.61 to −0.11)	0.015	0.00 (4; $p = 1.000$)	0	0.00 %	N/A
PC							
T_{pre} vs T_{post}	512	0.64 (0.06 to 1.22)	0.035	3.82 (10; $p = 0.955$)	0	0.00 %	1.21 ($p = 0.256$)
T_{post} vs T_{rec1}	542	−0.94 (−1.34 to −0.54)	<0.001	6.56 (10; $p = 0.766$)	0.221	0.00 %	1.26 ($p = 0.240$)
T_{post} vs T_{rec2}	461	−1.16 (−2.16 to −0.17)	0.027	7.26 (8; $p = 0.509$)	0	0.00 %	N/A
P-ACTH							
T_{pre} vs T_{post}	360	3.36 (1.66 to 5.07)	0.004	171.05 (5; $p < 0.0001$)	1.606	97.1 %	N/A
T_{post} vs T_{rec1}	360	−2.50 (−3.97 to −1.04)	0.007	132.39 (5; $p < 0.0001$)	1.439	96.2 %	N/A
HR							
T_{pre} vs T_{prep}	417	0.68 (0.47 to 0.89)	<0.0001	0.28 (11; $p = 1.000$)	0	0.00 %	1.39 ($p = 0.194$)
T_{pre} vs T_{speech}	510	1.13 (0.78 to 1.48)	<0.0001	0.61 (10; $p = 1.000$)	0	0.00 %	0.71 ($p = 0.499$)
T_{pre} vs T_{maths}	510	0.96 (0.65 to 1.26)	<0.0001	0.44 (10; $p = 1.000$)	0	0.00 %	1.28 ($p = 0.233$)
T_{prep} vs T_{rec}	428	−0.70 (−0.91 to −0.49)	<0.0001	0.22 (10; $p = 1.000$)	0	0.00 %	−1.20 ($p = 0.260$)
T_{speech} vs T_{rec}	409	−1.21 (−1.60 to −0.81)	<0.0001	0.77 (10; $p = 1.000$)	0	0.00 %	−0.56 ($p = 0.592$)
T_{maths} vs T_{rec}	389	−1.03 (−1.35 to −0.71)	<0.0001	1.49 (10; $p = 1.000$)	0	0.00 %	−1.00 ($p = 0.343$)

into four main categories: [i] T_{pre} = baseline, last timepoint before TSST; [ii] T_{post} = immediately (0–5 min) following TSST completion; [iii] T_{peak} = peak values after TSST (for SC with a response latency (mean = ~13 min, range = 10–30 min; this mean was derived from a *meta*-analysis study including a total of 34 papers (Liu et al., 2017 Aug); [iv] T_{rec1} = 30 min recovery after T_{post} (or T_{peak} if applicable), T_{rec2} = 60 min recovery after T_{post} (or T_{peak} if applicable). For certain physiological measures such as HR, time points during the task were additionally labelled as TSST preparation phase (T_{prep}), TSST speech phase (T_{speech}), and TSST maths phase (T_{maths}). Please refer to Supplement 2 Table S1C for further details about timepoint labelling.

2.3. Study quality assessment

Quality assessment was conducted by two independent raters (I.S.C. M. and R.S.). The risk of bias for each study was evaluated with the modified form from the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) (Sterne et al., 2016) (Supplementary Material 3). The quality rating for each study is shown in Supplementary Materials 2 Tables 2A-E & 3A-B.

2.4. Statistical analysis

All statistical analyses were performed with R statistical software (version 4.1.1) using packages '*meta*' and '*dmetar*' (Schwarzer et al., 2015). To synthesize the effect size, within-group standardized mean differences (SMD) and standard error (SE) were pre-calculated based on the means and standard deviations for each timepoint in each individual study. The SMD was equivalent to Cohen's *d* for repeated measures when the correlation between measures was accounted for (Lakens, 2013). The formula for calculation is as follows:

$$\text{Cohen}'\text{s } d = \frac{M_2 - M_1}{\sqrt{SD_1^2 + SD_2^2 - 2 \times r \times SD_1 \times SD_2}} \times \sqrt{2(1-r)}$$

$$\text{SE} = \frac{\sqrt{SD_1^2 + SD_2^2 - 2 \times r \times SD_1 \times SD_2}}{\sqrt{n}}$$

M_1 & M_2 = mean value of observation at timepoint 1 and timepoint 2 respectively.
SD_1 & SD_2 = standard deviation of observation at timepoint 1 and timepoint 2 respectively.
r = correlation between M_1 & M_2 (assumed to be 0.5).

The random-effect model was adopted to account for the heterogeneity across studies. Using the adjusted random-effects weights, the pooled effect size was calculated using the inverse variance method (i.e., individual studies' effect sizes were inversely weighted according to their respective standard error). The restricted maximum-likelihood estimator was used to calculate the heterogeneity variance (Viechtbauer, 2005). In order to reduce the probability of false positive outcomes, Hartung-Knapp adjustments were applied to control for the uncertainty in the estimate of heterogeneity, and to calculate the confidence interval (CI) around the pooled effect (Jackson et al., 2017; IntHout et al., 2014). The overall summary effect sizes were reported as SMD value with 95 % CI. The SMD effect sizes 0.2, 0.5 and 0.8 were interpreted as small, medium, and large respectively (Cohen, 2013).

Additionally, an outlier analysis test was implemented using the algorithm ('find.outliers'). Besides outlier identification, influence diagnostics were run to assess if the pooled effect was potentially distorted by some highly influential study (Viechtbauer and Cheung, 2010). This examination additionally considers studentized deleted residuals, DFFITS metric, Cook's Distance and covariance ratio to identify an influential case.

Forest plots were generated to display the pooled effect for each *meta*-analysis, containing the observed effect, CI, and weight of each study. Between-study heterogeneity was systematically evaluated using the I^2 statistics (Higgins and Thompson, 2002). The percentages 25 %, 50 % and 75 % represent low, moderate, and high heterogeneity respectively (Huedo-Medina et al., 2006). Funnel plots were generated for every analysis to illustrate potential publication bias. Egger's regression intercept test was used to inspect the asymmetry of the funnel plot if the analysis included at least 10 studies (Egger et al., 1997; Sterne et al., 2011). If the regression test revealed a potential publication bias ($p < 0.1$), the Duval & Tweedie trim and fill method was used to correct any biases towards extreme effect sizes (Duval and Tweedie, 2000).

If $N_{study} \geq 10$, *meta*-regression analyses were additionally performed

Table 2
meta-analysis of mood markers in response to TSST.

Mood markers (time contrasts)	n	SMD (95 % CI)	*p* value	Heterogeneity			Egger's test *t* stat (*p* value)
				Q stat (df; *p*)	τ^2	I^2	
Positive affect state							
T_{pre} vs T_{post}	332	−0.68 (−1.44 to 0.09)	0.076	46.82 (8; $p < 0.0001$)	0.784	82.9 %	N/A
T_{post} vs T_{rec}	222	1.17 (−0.31 to 2.65)	0.099	139.30 (5; $p < 0.0001$)	1.734	96.4 %	N/A
Negative affect state							
T_{pre} vs T_{post}	441	0.72 (0.61 to 0.83)	<0.0001	52.35 (12; $p < 0.0001$)	0.029	77.1 %	−2.12 ($p = 0.057$)
T_{post} vs T_{rec}	276	−1.04 (−1.39 to −0.69)	<0.001	260.28 (7; $p < 0.0001$)	0.182	97.3 %	N/A

to investigate potential modulating effects of relevant study-specific factors, including mean age of participants, publication year, TSST task duration and sex ratio (female-to-male), on main effect sizes (Higgins et al., 2019). The intercept of the model represented the pooled effect size after controlling for the modulating variables, which were demeaned before entering the model. The modulating effects were evaluated using *t* statistics and 95 % CI of the beta value.

3. Results

In total, 61 eligible studies were included in the systematic review, and 58 studies were included in the *meta*-analysis. Fig. 1 detailed the exclusion processes following the PRISMA guidelines.

3.1. Sample and study characteristics

The sample characteristics of the included studies are summarized in Supplementary Material S2 tables. The total sample size of healthy participants included in the *meta*-analyses was 2459 (1528 male and 931 female).[1] Sample size per study ranged from 10 to 114. Publication year spanned from 2005 to 2020. The age range was 18–65 years old (mean = 27.03, SD = 8.06). Most studies contained mixed male and female samples, with 10 studies having female-only participants and 12 studies having male-only participants.

3.2. Risk of bias and quality assessment

Supplementary Tables S2A-E and S3A-B list the total quality rating for all involved studies. The studies' ratings ranged from 15 to 25 (median value = 22). All studies' ratings fell in the first two quartiles, with 33/61 studies in the first quartile (≥22) and the remaining 28 in the second quartile (≥15).

3.3. Biomarkers

Time contrasts of each marker are included in Supplementary Tables (Tables S2A-S2E). The result summary of biomarkers and results of *meta*-regression model are described in Table 1 and Table 3 respectively. Details about outlier and publication bias for each analysis are included in Supplementary Material 4. All *meta*-analysis results below were obtained after outlier removal and adjusting for publication bias (where applicable).

(a) Salivary cortisol (SC)

See Table S2A in Supplementary Material 2 for study and time points details.

(i) T_{pre} vs T_{post}

Pooled effect sizes from 29 studies (N = 1050) revealed a significant increase of SC level at T_{post} compared to T_{pre} (SMD = 0.82, 95 %CI = [0.69, 0.95], $p < 0.0001$), suggesting anticipatory stress effect. Cross-study heterogeneity was high (I^2 95 %CI = [81.1 %, 89.7 %])). meta-regression analysis showed that the effect sizes of T_{post}-minus-T_{pre} remained moderately high and significant after controlling for the covariates (SMD = 0.70, $t(23) = 8.23$, $p < 0.0001$). Task length significantly moderated the effect size (b = 0.19, 95 %CI = [0.00, 0.39], $t(23) = 2.09$, $p < 0.05$), while publication year showed a marginal effect (b = 0.17, 95 %CI = [-0.01, 0.35], $t(23) = 1.99$, $p = 0.059$). Results are depicted in Supplementary Material 4 Fig. S1A.

(ii) T_{pre} vs T_{peak}

Pooled effect sizes from 28 studies (N = 1023) revealed a significant increase of SC level at T_{peak} (10–30 min after TSST) compared to T_{pre} (SMD = 0.85, 95 %CI = [0.73, 0.97], $p < 0.0001$), suggesting reactive stress effect. Cross-study heterogeneity was high (I^2 95 %CI = [85.1 %, 91.7 %]). meta-regression analysis showed that the effect sizes of T_{peak}-minus-T_{pre} remained moderately high and significant after controlling for the covariates (SMD = 0.82, $t(22) = 18.91$, $p < 0.0001$). All four variables were significant modulators of effect size: mean age (b = 0.10, 95 %CI = [0.02, 0.17], $t(22) = 18.91$, $p < 0.05$), publication year (b = -0.15, 95 %CI = [-0.17, −0.13], $t(22) = -13.83$, $p < 0.0001$), task length (b = -0.25, 95 %CI = [-0.38, −0.12], $t(22) = -4.07$, $p < 0.001$), and sex ratio (b = -0.23, 95 %CI = [-0.33, −0.14], $t(22) = -5.00$, $p < 0.0001$). Results are depicted in Supplementary Material 4 Fig. S1B.

(iii) T_{post} vs T_{rec1}

Pooled effect sizes from 21 studies (N = 852) revealed a significant decrease of SC level at T_{rec1}(30 min after TSST) compared to T_{post} (SMD = -0.31, 95 %CI = [-0.50, −0.13], $p < 0.01$). Cross-study heterogeneity was relatively low (I^2 95 %CI = [0.0 %, 57.3 %]). meta-regression analysis showed that the effect size of T_{rec}-minus-T_{post} was moderate and significant after controlling for the covariates (SMD = -0.47, $t(15) = -4.79$, $p < 0.001$). Task length was a significant modulator of effect size (b = 0.35, 95 %CI = [0.10, 0.60], $t(15) = 2.99$, $p < 0.01$). Results are depicted in Supplementary Material 4 Fig. S1C.

(iv) T_{post} vs T_{rec2}

Pooled effect sizes from 9 studies (N = 326) revealed a significant decrease of SC level at T_{rec2} (60 min after TSST) compared to T_{post} (SMD = -0.70, 95 %CI = [-0.89, −0.50], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 65.1 %]). Results are depicted in Supplementary Material 4 Fig. S1D.

(v) T_{peak} vs T_{rec1}

Pooled effect sizes from 20 studies (N = 712) revealed a significant decrease of SC level at T_{rec1} (30 min recovery after TSST) compared to T_{peak} (SMD = -0.98, 95 %CI = [-1.04, −0.91], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 47.2 %]). meta-regression

[1] One study (Garcia-Rubio et al., 2017) was not included in the sample size calculation due to unknown sex distribution (N = 21).

Table 3
Multiple *meta*-regression model of biomarkers and mood markers.

Markers (time contrasts)	Estimate	SE	*t-value*	95% CI	*p value*
Modulators					
Salivary cortisol (Tpre vs Tpost)					
Intercept	0.7	0.09	8.23	0.53 to 0.88	<.0001
Mean age of participants	0	0.07	-0.04	-0.15 to 0.14	0.972
Publication year	0.17	0.09	1.99	-0.01 to 0.35	0.059
Task length	0.19	0.09	2.09	0.00 to 0.39	0.048
Sex ratio (F-to-M)	-0.01	0.06	-0.2	-0.14 to 0.11	0.844
Salivary cortisol (Tpre vs Tpeak)					
Intercept	0.82	0.04	18.91	0.73 to 0.91	<.0001
Mean age of participants	0.1	0.04	2.68	0.02 to 0.17	0.014
Publication year	-0.15	0.01	-13.83	-0.17 to -0.13	<.0001
Task length	-0.25	0.06	-4.07	-0.38 to -0.12	<.001
Sex ratio (F-to-M)	-0.23	0.05	-5	-0.33 to -0.14	<.0001
Salivary cortisol (Tpost vs Trec1)					
Intercept	-0.47	0.1	-4.79	-0.68 to -0.26	<.001
Mean age of participants	0.11	0.07	1.45	-0.05 to 0.27	0.169
Publication year	0.17	0.11	1.51	-0.07 to 0.40	0.153
Task length	0.35	0.12	2.99	0.10 to 0.60	0.009
Sex ratio (F-to-M)	-0.05	0.09	-0.55	-0.24 to 0.14	0.59
Salivary cortisol (Tpeak vs Trec1)					
Intercept	-0.96	0.05	-16.76	-1.08 to -0.83	<.0001
Mean age of participants	-0.04	0.05	-0.72	-0.15 to 0.08	0.484
Publication year	-0.12	0.14	-0.89	-0.42 to 0.18	0.39
Task length	-0.1	0.08	-1.26	-0.27 to 0.07	0.23
Sex ratio (F-to-M)	0.05	0.08	0.59	-0.13 to 0.22	0.568
Plasma cortisol (Tpre vs Tpeak)					
Intercept	3.01	1.96	1.54	-1.79 to 7.81	0.176
Mean age of participants	2.44	5.22	0.47	-10.32 to 15.21	0.656
Publication year	2.21	0.78	2.86	0.32 to 4.11	0.029
Task length	2.52	3.02	0.84	-4.86 to 9.90	0.436
Sex ratio (F-to-M)	0.2	3.57	0.05	-8.54 to 8.93	0.958
Plasma cortisol (Tpost vs Trec1)					
Intercept	-1.01	0.33	-3.07	-1.82 to -0.21	0.022
Mean age of participants	-0.72	0.94	-0.77	-3.02 to 1.57	0.469
Publication year	-0.36	0.54	-0.66	-1.69 to 0.98	0.536
Task length	-0.6	0.65	-0.93	-2.18 to 0.98	0.387
Sex ratio (F-to-M)	-0.46	0.89	-0.52	-2.62 to 1.71	0.624
Heart rate (Tpre vs Tprep)					
Intercept	0.68	0.12	5.5	0.38 to 0.98	0.002
Mean age of participants	-0.08	0.18	-0.41	-0.52 to 0.37	0.694
Publication year	-0.15	0.19	-0.77	-0.61 to 0.32	0.47
Task length	-0.17	0.17	-1	-0.58 to 0.24	0.356
Sex ratio (F-to-M)	0.11	0.15	0.74	-0.26 to 0.49	0.487
Heart rate (Tpre vs Tspeech)					
Intercept	1.05	0.16	6.44	0.63 to 1.47	0.001
Mean age of participants	0.11	0.21	0.51	-0.44 to 0.66	0.635
Publication year	0.16	0.36	0.46	-0.75 to 1.08	0.663
Task length	0.26	0.38	0.68	-0.72 to 1.25	0.527
Sex ratio (F-to-M)	0.17	0.26	0.64	-0.49 to 0.83	0.548
Heart rate (Tpre vs Tmaths)					
Intercept	0.93	0.16	5.86	0.52 to 1.34	0.002
Mean age of participants	0.02	0.23	0.11	-0.56 to 0.60	0.92
Publication year	0.02	0.35	0.07	-0.87 to 0.91	0.947
Task length	0.08	0.38	0.2	-0.91 to 1.06	0.85
Sex ratio (F-to-M)	0.19	0.25	0.79	-0.44 to 0.83	0.468
Heart rate (Tprep vs Trec)					
Intercept	-0.76	0.12	-6.48	-1.06 to -0.46	0.001
Mean age of participants	-0.27	0.25	-1.09	-0.92 to 0.37	0.325
Publication year	0.07	0.19	0.36	-0.41 to 0.54	0.73
Task length	0.09	0.17	0.54	-0.35 to 0.53	0.612
Sex ratio (F-to-M)	0.04	0.18	0.25	-0.42 to 0.51	0.812
Heart rate (Tspeech vs Trec)					
Intercept	-1.14	0.15	-7.48	-1.54 to -0.75	<0.001
Mean age of participants	-0.26	0.21	-1.26	-0.79 to 0.27	0.264
Publication year	-0.31	0.33	-0.92	-1.17 to 0.55	0.4
Task length	-0.43	0.36	-1.19	-1.36 to 0.50	0.286
Sex ratio (F-to-M)	-0.14	0.24	-0.59	-0.76 to 0.48	0.58
Heart rate (Tmaths vs Trec)					
Intercept	-1.03	0.14	-7.32	-1.39 to -0.67	<0.001
Mean age of participants	-0.17	0.2	-0.84	-0.69 to 0.35	0.437
Publication year	-0.16	0.31	-0.53	-0.95 to 0.63	0.62
Task length	-0.25	0.34	-0.73	-1.12 to 0.62	0.499
Sex ratio (F-to-M)	-0.18	0.22	-0.83	-0.75 to 0.38	0.445
Negative affect state (Tpre vs Tpost)					
Intercept	0.69	0.06	10.8	0.54 to 0.84	<.0001
Mean age of participants	0.04	0.13	0.32	-0.26 to 0.34	0.757
Publication year	0.09	0.07	1.2	-0.08 to 0.25	0.268
Task length	0.14	0.11	1.27	-0.12 to 0.39	0.245
Sex ratio (F-to-M)	-0.01	0.08	-0.16	-0.21 to 0.18	0.875

analysis showed that the effect size of T_{rec1}-minus-T_{peak} remained high and significant after controlling for the covariates (SMD = -0.96, $t(14)$ = -16.76, $p < 0.0001$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S1E.

(vi) T_{peak} vs T_{rec2}

Pooled effect sizes from 8 studies (N = 303) revealed a significant decrease of SC level at T_{rec2} (60 min recovery after TSST) compared to T_{peak} (SMD = -0.57, 95 %CI = [-0.72, −0.41], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 67.6 %]). Results are depicted in Supplementary Material 4 Fig. S1F.

(b) Salivary alpha-amylase (sAA).

See Table S2B in Supplementary Material 2 for study details and time points included in *meta*-analyses of sAA.

(i) T_{pre} vs T_{post}

Pooled effect sizes from 9 studies (N = 314) revealed a significant increase of sAA level at T_{post} compared to T_{pre} (SMD = 1.18, 95 %CI = [1.10, 1.26], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 64.8 %])). Results are depicted in Supplementary Material 4 Fig. S2A.

(ii) T_{post} vs T_{rec1}

Pooled effect sizes from 5 studies (N = 199) revealed a significant decrease of sAA level at T_{rec1} compared to T_{post} (SMD = -0.36, 95 %CI = [-0.61, −0.11], $p < 0.05$). Cross-study heterogeneity was moderate (I^2 95 %CI = [0.0 %, 79.2 %]). Results are depicted in Supplementary Material 4 Fig. S2B.

(iii) T_{post} vs T_{rec2}

No *meta*-analysis due to insufficient study number ($N_{study} < 5$).

(c) Plasma cortisol (PC).

See Table S2C in Supplementary Material 2 for study and time points details.

(a) T_{pre} vs T_{post}

Pooled effect sizes from 11 studies (N = 512) revealed a significant increase of PC level at T_{post} compared to T_{pre} (SMD = 0.64, 95 %CI = [0.06, 1.22], $p < 0.05$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 60.2 %]). meta-regression analysis revealed that the pooled effect size showed insignificant difference after controlling for the covariates (SMD = 3.01, $t(6) = 1.54$, $p = 0.18$). Publication year was a significant modulator of effect size (b = 2.21, 95 %CI = [0.32, 4.11], $t(6) = 2.86$, $p < 0.05$). Results are depicted in Supplementary Material 4 Fig. S3A.

(b) T_{post} vs T_{rec1}

Pooled effect sizes from 11 studies (N = 542) revealed a significant decrease of PC level at T_{rec1} compared to T_{post} (SMD = -0.94, 95 %CI = [-1.34, −0.54], p < 0.001). Cross-study heterogeneity was low (I^2 95 % CI = [0.0 %, 60.2 %]). meta-regression analysis showed that the effect size remained large and significant after controlling for the covariates (SMD = -1.01, $t(6) = -3.07$, $p < 0.05$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S3B.

(c) T_{post} vs T_{rec2}

Pooled effect sizes from 9 studies (N = 391) revealed a significant decrease of PC level at T_{rec2} compared to T_{post} (SMD = -1.16, 95 %CI = [-2.16, −0.17], $p < 0.05$. Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 64.8 %]). Results are depicted in Supplementary Material 4 Fig. S3C.

(d) Plasma adrenocorticotropic hormone (P-ACTH).

See Table S2D in Supplementary Material 2 for study details and time points included in *meta*-analyses of P-ACTH.

(i) T_{pre} vs T_{post}

Pooled effect sizes from 6 studies (N = 360) revealed a significant increase of P-ACTH levels at T_{post} compared to T_{pre} (SMD = 3.36, 95 % CI = [1.66, 5.07], $p < 0.01$). Cross-study heterogeneity was substantially high (I^2 95 %CI = [95.4 %, 98.1 %]). Results are depicted in Supplementary Material 4 Fig. S4A.

(ii) T_{post} vs T_{rec1}

Pooled effect sizes from 6 studies (N = 360) revealed a significant decrease of P-ACTH level at T_{rec1} compared to T_{post} (SMD = -2.50, 95 % CI = [-3.97, −1.03], $p < 0.01$). Cross-study heterogeneity was high (I^2 95 %CI = [93.9 %, 97.7 %]). Results are depicted in Supplementary Material 4 Fig. S4B.

(iii) T_{post} vs T_{rec2}

No *meta*-analysis due to insufficient study number ($N_{study} < 5$).

(e) Heart rate (HR)

See Table S2E in Supplementary Material 2 for study details and time points included in *meta*-analyses of HR.

(i) T_{pre} vs T_{prep}

Pooled effect sizes from 12 studies (N = 417) revealed a significant increase of HR at T_{prep} compared to T_{pre} (SMD = 0.68, 95 %CI = [0.47, 0.89], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 58.3 %])). meta-regression analysis revealed that the effect size remained significant after controlling for the covariates (SMD = 0.68, $t(6) = 5.50$, $p < 0.01$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S5A.

(ii) T_{pre} vs T_{speech}

Pooled effect sizes from 11 studies (N = 510) revealed a significant increase of HR at T_{speech} compared to T_{pre} (SMD = 1.13, 95 %CI = [0.78, 1.48], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 60.2 %]). meta-regression analysis revealed that the effect size remained significant after controlling for the covariates (SMD = 1.05, $t(5) = 6.44$, $p < 0.01$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S5B.

(iii) T_{pre} vs T_{maths}

Pooled effect sizes from 11 studies (N = 510) revealed a significant increase of HR at T_{maths} compared to T_{pre} (SMD = 0.96, 95 %CI = [0.65, 1.26], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 60.2 %]). meta-regression analysis revealed that the effect size remained significant after controlling for the covariates (SMD = 0.93, $t(5) = 5.86$, $p < 0.01$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S5C.

(iv) T_{prep} vs T_{rec}

Pooled effect sizes from 11 studies (N = 428) revealed a significant decrease of HR levels at T_{rec} compared to T_{prep} (SMD = -0.70, 95 %CI = [-0.91, −0.49], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 % CI = [0.0 %, 60.2 %])). meta-regression analysis revealed that the effect size remained significant after controlling for the covariates (SMD = -0.76, $t(5) = -6.48$, $p < 0.01$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S5D.

(v) T_{speech} vs T_{rec}

Pooled effect sizes from 11 studies (N = 409) revealed a significant decrease of HR at T_{rec} compared to T_{speech} (SMD = -1.21, 95 %CI = [-1.60, −0.81], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 % CI = [0.0 %, 60.2 %])). meta-regression analysis revealed that the effect size remained large and significant after controlling for the covariates (SMD = -1.14, $t(5) = -7.50$, $p < 0.001$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S5E.

(vi) T_{maths} vs T_{rec}

Pooled effect sizes from 13 studies (N = 389) revealed a significant decrease of HR T_{rec} compared to T_{maths} (SMD = -1.03, 95 %CI = [-1.35, −0.71], $p < 0.0001$). Cross-study heterogeneity was low (I^2 95 %CI = [0.0 %, 60.2 %]). meta-regression analysis revealed that the effect size remained large and significant after controlling for the covariates (SMD = -1.03, $t(5) = -7.32$, $p < 0.001$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S5F.

(f) Descriptive summaries of biomarkers.

Due to the small study number available, the effects on these markers are systematically reviewed as below.

Salivary biomarkers

Dehydroepiandrosterone (DHEA): Two studies (N = 33 and 40, predominantly males) found significant increase in salivary DHEA levels immediately post-TSST (Izawa et al., 2008), or at 10 and 20 min after TSST (Sugaya et al., 2012), compared to pre-TSST level. However, in

these studies, while 70 % of participants exhibited significant and stable increase in DHEA level, the remaining participants showed large variations and fluctuations in response. Therefore, group-averaged DHEA responses should be interpreted with caution due to substantial heterogeneity among individuals. Izawa et al. (Izawa et al., 2008) additionally investigated DHEA-sulfate (DHEA-S) level and the DHEA-S/DHEA ratio, but found no significant difference in either measure before and after TSST.

Salivary interleukin-6 (IL-6): Two studies (N = 34 and 50, predominantly males) reported elevated IL-6 levels in response to TSST (Janusek et al., 2017; Izawa et al., 2013). Results from Izawa and colleagues (Izawa et al., 2013) suggested that IL-6 levels were increased during TSST and remained elevated for 20 min. Janusek and colleagues (Janusek et al., 2017) also found increase in IL-6 levels, followed by prolonged elevation beyond termination of stressor. Furthermore, Izawa et al. (Izawa et al., 2013) found persistence increase in IL-6 levels at 60 min following TSST in some, but not all, individuals. Thus, the recovery trajectory of salivary IL-6 is still unclear and may vary substantially across individuals.

Salivary C-reactive protein (CRP): Two studies (N = 39 and 15, all females and predominantly males respectively) (Kimura et al., 2013; Kennedy et al., 2014) showed increases in salivary CRP levels in response to TSST. Kimura and colleagues (Kimura et al., 2013) further divided participants into cortisol responders and non-responders and found a significant elevation of salivary CRP levels after the TSST preparation and speech phases compared to pre-TSST level in responders, but not in non-responders.

Blood biomarkers

Plasma interleukin-6 (IL-6): Four studies (N = 18–69, mostly mixed-sex (Christian et al., 2013; Böbel et al., 2018; Carpenter et al., 2010; Yamakawa et al., 2015) consistently found plasma IL-6 levels increase following TSST. Among these, two studies found increase in IL-6 level that persisted until at least 120 min after TSST in participants living in urban area (Böbel et al., 2018), and in white or African American women (Christian et al., 2013). However, Böbel et al. (Böbel et al., 2018) reported that the plasma IL-6 levels peaked at 90 min after TSST and returned to baseline at 120 min post-TSST among participants living in rural areas. Furthermore, Carpenter et al. (Carpenter et al., 2010) found that the plasma IL-6 levels increased persistently for at least 60 min, with greater increase in individuals with childhood maltreatment. Lastly, Yamakawa et al. (Yamakawa et al., 2015) analyzed the IL-6/IL-10 ratio, a parameter representing the balance of pro- and anti-inflammation. They did not find significant change of ratio after task compared to baseline, though a trend of increase was observed from the baseline to at least 30 min after task, and the change differed between subgroups possessing different polymorphisms of the serotonin transporter gene.

Prolactin: Four studies (N = 23–98, mostly males) investigated plasma prolactin levels before and after TSST (Chong et al., 2008; Klumbies et al., 2014; Munro et al., 2005; Uhart et al., 2006) and reported mixed results. An increase of plasma prolactin levels was found by Chong et al. (Chong et al., 2008) and Klumbies et al. (Klumbies et al., 2014) after the TSST compared to baseline. However, Uhart et al.'s results indicate that the increase in prolactin levels was specific to Caucasian American, but not in African American (Uhart et al., 2006). Finally, Munro et al. (Munro et al., 2005) found no change in prolactin levels after task compared to baseline, and concluded that prolactin may not be a valid stress hormone to assess TSST response. The differential prolactin responses could be partly due to difference in sample size and sex ratio across studies.

Other physiological markers

Skin conductance (SCL): Five studies (N = 15–58, predominantly females) (Kennedy et al., 2014; Cărnuţă et al., 2015; Hendrawan et al., 2012; Aleknaviciute et al., 2016; Wearne et al., 2019) consistently indicated elevated SCL values during TSST phases compared to baseline. However, the recovery pattern of SCL levels after task showed inconsistent results. Two studies based on all-female samples found significant decrease of SCL level after recovery phases to a level comparable to pre-TSST level, either after 45-minutes recovery (Aleknaviciute et al., 2016), or after 15-minute recovery (Cărnuţă et al., 2015). In contrast, Hendrawan et al. (Hendrawan et al., 2012) and Wearne et al.'s (Wearne et al., 2019) results based on all-male and mixed-sex samples indicated that despite reductions of SCL levels after 20-minute recovery, the levels after recovery remained significantly elevated compared to baseline. A similar pattern was observed in Kennedy et al.'s study (Kennedy et al., 2014) including 15 females. Thus, the recovery trajectory of SCL levels remains unclear.

Blood pressure (BP): Three studies (N = 30–40, mostly males) reported BP measures in response to TSST (Izawa et al., 2008; Böbel et al., 2018; Chen et al., 2017). Two studies reported elevation of systolic blood pressure (SBP) during the TSST compared to baseline (Izawa et al., 2008; Chen et al., 2017), whereas significant increase of diastolic blood pressure (DBP) during TSST was found only in the latter study based on all-male sample (Izawa et al., 2008), but not in the former study based on mixed-sex sample (Chen et al., 2017). Böbel et al. (Böbel et al., 2018) assessed the median arterial pressure, calculated according to the formula: DBP + (SBP - DBP)/3. This measure showed an increase during task compared with basal values.

Heart rate variability (HRV): Five studies (N = 18–78, mostly mixed-sex) examined HRV in time or frequency domain (Izawa et al., 2013; Yamakawa et al., 2015; Klumbies et al., 2014; Wearne et al., 2019; Lackschewitz et al., 2008). Three studies showed decrease in HRV in the time domain during the TSST preparation, speech and maths phases compared to baseline (Izawa et al., 2013; Klumbies et al., 2014; Lackschewitz et al., 2008). Lackschewitz et al. (Lackschewitz et al., 2008) additionally revealed recovery of HRV in the time domain at 15 min after TSST from the immediately post-TSST level. For HRV in the low frequency (LF) domain, studies showed mixed findings, with one study showing significant difference from baseline to TSST anticipation phase (Lackschewitz et al., 2008), while another study showing no such change (Wearne et al., 2019). However, both studies reported increase of low-frequency HRV during recovery compared to during TSST. For high-frequency (HF) HRV, two studies found significant reductions during TSST compared to baseline, which fully recovered at 15–20 min after TSST completion (Wearne et al., 2019; Lackschewitz et al., 2008). Two studies further examined the LF/HF HRV ratio during TSST (Yamakawa et al., 2015; Lackschewitz et al., 2008). The latter study based on a small all-male sample did not find any significant change of the LF/HF HRV before and during TSST, while the former study based on a small mixed-sex sample showed significant decrease of LF/HF ratio after TSST completion compared to during TSST.

3.4. Emotional state measures

The time contrasts of each marker are detailed in Supplementary Material 2 Tables S3A-3B. The result summary of mood measures and results of *meta*-regression model are described in Table 2 and Table 3 respectively. Details about outlier and publication bias for each analysis are included in Supplement 4. All *meta*-analysis results below were obtained after outlier removal and adjusting for publication bias (where applicable).

(a) Positive affect (PA) state

See Table S3A in Supplementary Material 2 for study details and time points included in *meta*-analyses of PA.

(i) T_{pre} vs T_{post}

Pooled effect sizes from 9 studies (N = 332) revealed marginal decrease of PA state levels at T_{post} compared to T_{pre} (SMD = -0.68, 95 % CI = [-1.44, 0.09], $p = 0.08$). Cross-study heterogeneity was high (I^2 95 %CI = [69.0 %, 90.6 %]). Results are depicted in Supplementary

Material 4 Fig. S6A.

(ii) T_{post} vs T_{rec}

Pooled effect sizes from 6 studies (N = 222) revealed marginal increase of PA state at T_{rec} compared to T_{post} (SMD = 1.17, 95 %CI = [-0.31, 2.65], p = 0.10). Cross-study heterogeneity was substantially high (I^2 95 % CI = [94.2 %, 97.8 %]). Results are depicted in Supplementary Material 4 Fig. S6B.

(b) Negative affect (NA) state

See Table S3B in Supplementary Material 2 for study details and time points included in *meta*-analyses of NA.

(i) T_{pre} vs T_{post}

Pooled effect sizes from 13 studies (N = 441) revealed a significant increase of NA state at T_{post} compared to T_{pre} (SMD = 0.72, 95 %CI = [0.61, 0.83], $p < 0.0001$). The cross-study heterogeneity was high (I^2 95 % CI = [61.0 %, 86.5 %]). meta-regression analysis showed that the effect size remained moderate and significant after controlling for the covariates (SMD = 0.69, $t(7)$ = 10.80, $p < 0.0001$). No significant modulating effect was found ($p > 0.1$). Results are depicted in Supplementary Material 4 Fig. S7A.

(ii) T_{post} vs T_{rec}

Pooled effect sizes from 8 studies (N = 276) revealed a significant decrease of NA state at T_{rec} compared to T_{post} (SMD = -1.04, 95 %CI = [-1.39, −0.69], $p < 0.001$). Heterogeneity level was substantially high (I^2 95 % CI = [96.1 %, 98.1 %]). Results are depicted in Supplementary Material 4 Fig. S7B.

3.5. Correlations between biological and mood markers to TSST

We additionally explored the association between the biological and emotional response changes to TSST. Due to insufficient study number on the same response measures, these studies are reviewed systematically below. Table S4A in Supplementary Material 2 presents a descriptive summary of the correlation analysis findings.

In total, 6 studies reported the correlations between changes of biomarkers and mood measures during and after TSST (Petrakova et al., 2015; Cărnuţă et al., 2015; Monteleone et al., 2018; Edelstein et al., 2010; Rimmele et al., 2007; Zhang et al., 2019). In general, the study results were inconsistent. One study including a female-only sample (N = 62) reported significantly negative correlation between SC response to the TSST and state anxiety changes after TSST compared to baseline (Cărnuţă et al., 2015). However, another study using a mixed-sex sample (N = 50) found no significant correlation between SC response to the TSST and PA or NA changes after TSST (Zhang et al., 2019). In addition, one study including 44 males found significantly positive correlation between HR change to TSST and calmness rating change before and after TSST (Rimmele et al., 2007), while Zhang et al. (Zhang et al., 2019) found no significant correlation between HR responses and PA or NA changes following TSST. For studies additionally including the recovery phase, Rimmele et al. (Rimmele et al., 2007) found significant positive correlation between SC and calmness rating changes, and significantly negative correlation between SC and state anxiety rating changes, throughout the entire course of TSST. However, two other studies based on mixed-sex or female-only samples failed to find any significant correlation between SC changes and state positive, state negative or state anxiety rating changes (Monteleone et al., 2018; Edelstein et al., 2010). Moreover, changes in other physiological markers including HR, SCL and sAA throughout the TSST response and recovery phases showed no significant correlation with state anxiety (HR, SCL) or NA state (sAA) changes (Petrakova et al., 2015; Cărnuţă et al., 2015).

4. Discussion

In this *meta*-analysis, we found that salivary cortisol was the most reliable biomarker which showed increase during acute stress and decrease after recovery. Other biomarkers such as HR, sAA and P-ACTH also showed statistically significant increase and decrease during acute stress and recovery, but the available number of studies was considerably smaller. Systematic reviews indicate that biomarkers such as DHEA, IL-6 and CRP, and SCL levels received preliminary supporting evidence for their sensitivity to acute stress, but their response profiles during stress and following recovery may show considerable individual differences. Heterogeneity in findings across studies could be partly due to participants' past life stress experience. In terms of psychological measures, general negative affect showed significant increase to acute stress, and decrease following recovery, although the numbers of available studies were modest (~10). In contrast, positive affect showed only marginal changes during and after TSST.

Supplementary analyses on postmenopausal women were generally consistent with those on younger participants. Modest positive evidence was obtained on cortisol level increase and decrease during response to and recovery from the TSST respectively, although negative findings also existed which could be due to insufficient statistical power. Similar responses to and recovery from the TSST were observed for both HR and BP (both systolic and diastolic). Increases and decreases in negative affects were also observed among postmenopausal women during response to and recovery from the TSST, provided the study recruited sufficient sample for higher statistical power.

Generally, relatively fewer studies explicitly assessed the between-participant interrelations between biological and emotional response *changes* across key timepoints (e.g., peak-minus-pre-task, recovered-minus-peak) during and following TSST, and the findings were mixed. Between-study differences in sample size, nature of measures and analysis methods may contribute to the heterogeneity in findings.

4.1. HPA axis

The activation of HPA axis in response to acute stress provides the mechanistic basis for the observed differential stress reactivity and recovery profiles in cortisol and ACTH. The HPA axis plays an important role in regulating the production and secretion of glucocorticoid hormones (Beurel and Nemeroff, 2014; Fulford et al., 2005; Spiga and Lightman, 2015). Following HPA activation, the androgen precursors DHEA and its sulphated metabolite DHEA-S are also secreted by the adrenal cortex. Previous *meta*-analyses have examined the modulatory effects of sex (Liu et al., 2017), protocol (Goodman et al., 2017) and age (Seddon et al., 2020) on cortisol responses to the TSST, while the present *meta*-analysis results supported the sensitivity and reliability of SC responses to both TSST and recovery across studies. On the other hand, although DHEA secretion shares similar temporal trajectory as cortisol, the present *meta*-analysis results were too limited to indicate the sensitivity of DHEA to TSST-induced acute stress. Also, our *meta*-analysis results did not show significant PC responses to TSST. While SC reflects the levels of biologically active unbound cortisol, PC reflects the total cortisol level including both protein-bound and free cortisol (Dickerson and Kemeny, 2004). Thus, the results suggest that unbound free cortisol may be particularly sensitive to TSST-induced acute stress.

Meta-regression results suggested that publication year, mean age of participants, TSST task length and sex ratio all significantly modulated SC response to reactive stress. The included studies were published between 2005 and 2020. It could be that cortisol testing and analysis methods advanced in recent years, particularly in the previous decade (van den Ouweland and Kema, 2012). An earlier *meta*-analysis also reported significant modulating effect of age, and additionally found that females showed almost three times stronger effect of age than males

(Otte et al., 2005). The authors considered that the sex effect may be attributed to sex-specific change of reproductive hormones with age. In addition, our *meta*-analysis revealed task length as a significant modulator of SC response to anticipatory stress and its recovery, despite previous mixed results on the effect of TSST duration on acute stress reactivity (e.g., Dickerson and Kemeny, 2004; Goodwin et al., 2017).

4.2. SAM axis

The SAM axis is also triggered by acute stress to exert control over the visceral activities and glandular functions of the body through the autonomic nervous system and the endocrine system. For instance, HR is regulated via homeostasis between sympathetic and parasympathetic (vagal) activities. Upon exposure to a stressor, the sympathetic nervous system is activated and the parasympathetic system is suppressed, resulting in increased HR. Our *meta*-analysis results showed reliable HR increases during acute stress compared to baseline, and decreases during recovery from stress.

At the same time of sympathetic activation, under the regulation of baroreflexes and the autonomic effector systems, an elevation of total peripheral resistance of blood vessels and vasoconstriction results in increased arterial pressure (Joyner et al., 2008). Based on the 2 studies included in our systematic review, SBP appeared more consistently sensitive to acute stress compared to DBP, particularly for females. However, the number of studies was too low to reach meaningful conclusion. For the same reason, it is also unclear whether other measures of BP, such as mean arterial pressure, may be reliable markers of acute stress reactivity.

Furthermore, sympathovagal balance characterizes the autonomic state as a result of the sympathetic and vagal influences (Goldberger, 1999), which is often indexed by HRV in time or frequency domain. Due to the limited studies with the same HRV measure, the results remain inconclusive, although the general pattern was that time- and frequency-domain HRV decreased during TSST, and increased following recovery. SCL, on the other hand, is primarily under the sympathetic control and reflects a state of arousal (Sequeira et al., 2009). We found the SCL showed typical stress reactivity during TSST, but inconsistent recovery curve, with some studies reporting extended elevation of SCL levels. The distinct patterns of SCL and HR recovery may be due to differential contributions of sympathetic and parasympathetic systems during acute stress response (Storm et al., 2002).

sAA is another candidate marker of SAM activity (Nater and Rohleder, 2009), with a main role in digestive function. The secretion of sAA is stimulated by the sympathetic nervous system activation and inhibited by parasympathetic activation (Anderson et al., 1984). Our current *meta*-analysis results showed a significant increase in sAA levels in response to TSST and a decrease during recovery. This result is consistent with documented sAA increase in response to other types of acute stress, such as watching stressful video (Takai et al., 2004; Bosch et al., 2003) and performing memory or mental arithmetic task (Noto et al., 2005; Edwards et al., 2006). These findings collectively support the global sensitivity of sAA to acute stress.

4.3. Immunological system

Two widely studied immune markers, IL-6 and CRP, showed stress reactivity responses. However, the elevation was prolonged, and there was no observable recovery at 30 min post-task. Previous research finding indicated that the IL-6 response trajectory was different in men and women, with earlier peak at around 30 min post-task for men, but a later peak at >60 min after task for women (Khera et al., 2005). For CRP response, significant effects of sex and race were also found (Coelho et al., 2014). Therefore, delineation of IL-6 and CRP response and recovery to acute stress may need to be separately performed for different sex and racial groups. Moreover, childhood maltreatment may enhance and prolong the inflammatory responses to acute stress (Fomicheva et al., 2004), thus future studies assessing IL-6 and CRP responses would need to account for prior major stress exposure.

Prolactin has also been reported as a regulator of stress response. Though the underlying mechanism remains unclear, it has been proposed that the protective role of prolactin against stress damage is through immunoenhancement (Lennartsson and Jonsdottir, 2011). Previous evidence indicates the levels of prolactin response was associated with the HPA axis stress responses (Purves et al., 2001). However, the current limited studies did not provide evidence showing prolactin as a reliable marker of acute stress response.

4.4. Emotion system and Emotion-biology association

We also assessed changes of emotional state throughout TSST. According to the classic James-Lange theory, emotions emerge from physiological changes, and are closely linked to visceral functions (Lemaire et al., 2011). Also, brain systems for emotion and physiological stress responses highly overlap and interact, underpinning the possibility that they should be interrelated (Dickerson et al., 2004; Sloman et al., 2003; Slavich and Irwin, 2014; Pizzagalli, 2014). Results from our *meta*-analysis found reliable increase of negative emotions to TSST, which also showed reliable decreases after recovery. These results aligned with the general view that the social evaluative threat induced by TSST signals defeat and diminished social status (Shacham, 1983), which were linked to high negative emotion and major depressive disorder (MDD) (Watson et al., 1988). On the other hand, our results revealed only marginally significant increase and decrease of PA level in response to TSST and following recovery. Previous studies revealed a strong association between stress and MDD (Whittle et al., 2011). While chronic stress exposure could have major impacts on the brain reward system (Nolen-Hoeksema and Aldao, 2011), our current results suggest that acute stress responses may impact more on negative rather than positive mood state. The intricate interaction of emotional and biological responses to acute and chronic stress, and their roles in precipitating emotion dysregulations and disorders, would be a topic of great interest for future translational and clinical studies.

We also found considerable heterogeneity in the findings on state affect. One source of heterogeneity may be the different questionnaires used to measure the emotions. For example, NA state was measured using POMS (Golatowski et al., 2013), the NA subscale of the Positive and Negative Affect Schedule (PANAS) (Shirtcliff et al., 2001), or study-specific VAS. While both POMS and PANAS are well-established state affect questionnaires, their measurements may not necessarily converge, and considerable variations can exist in the study-specific VAS measures. Another potential important cause of heterogeneity is the different sex composition across studies, given known sex difference in emotion reactivity and regulation (Gallagher et al., 2006). Age could also be an important factor impacting emotion regulation (Raff et al., 2002).

We additionally reviewed correlations between the state changes in the biological and affective domains. During both TSST and recovery, significant negative associations were found between SC and state anxiety changes (Cărnuţă et al., 2015; Rimmele et al., 2007). One study also found SC changes correlated positively with calmness rating throughout the entire course of TSST (Rimmele et al., 2007). However, other studies failed to obtain significant relationship between SC and state anxiety changes, and between SC and changes in PA or NA (Monteleone et al., 2018; Edelstein et al., 2010; Zhang et al., 2019). Moreover, although one study revealed positive correlation between HR and calmness rating to TSST in male participants (Rimmele et al., 2007), another study including mixed-sex participants revealed no correlation between HR and PA or NA change (Zhang et al., 2019). Several sources of between-study heterogeneity may account for the discrepant findings, including the method of calculating physiological and affective changes during task and recovery, the scales used to assess state affect, sample size and characteristics, and potentially insufficient statistical power for

some mixed-sex samples with modest sample size.

It is worth further discussing the potential effects of different ways of measuring the key biomarkers during and after the TSST. For salivary cortisol, the sampling methods across studies were mainly synthetic swab or passive drooling (Bae et al., 2016). Previous studies tended to suggest high consistency of cortisol readings across both sampling methods, although DHEA measurement may vary to greater extents depending on the method of collection (Ferreira et al., 2021; Koray et al., 2003). Regarding the quantification methods of cortisol concentration, the common methods include enzyme-linked immunoassay (EIA), radioimmunoassay (RIA) and liquid chromatography-tandem mass spectrometry (LC-MS/MS). While the EIA method may return higher absolute cortisol level readings than RIA, the change of cortisol levels across timepoints should be relatively unaffected (since the same quantification method was applied on samples across all timepoints) (Merswolken et al., 2013). Furthermore, the immunoassay and LC-MS/MS methods were found to show largely comparable results on salivary cortisol response dynamics (Kische et al., 2021). For HR, the measurement methods were mainly electrocardiogram (ECG) and heart rate monitors (HRM). A recent review comprehensively compared the HR measurement using different methods, and found all methods including plethysmographic, ballistic and electrical showed excellent estimation for HR in steady laboratory conditions, such as during a TSST paradigm which involves little physical activity (Dedovic et al., 2009).

Although several previous studies associated changes in SC and affect-related measures, these studies focused on basal or morning-awakening cortisol responses and relatively long-term anxiety measures (e.g., Koray et al., 2003; Merswolken et al., 2013; Kische et al., 2021; Dedovic et al., 2009). The current study suggested that SC state changes during TSST may be particularly linked with changes in state anxiety (or lack of anxiety, namely calmness). Associations between other physiological and mood changes to TSST or during recovery were yet to be established. These results are consistent with known strong modulatory effect of cortisol on important brain regions involved in emotion processing and regulation, such as the amygdala, hippocampus and prefrontal cortex (Dedovic et al., 2009; Harrewijn et al., 2020). Our findings thus highlight the intimate interactions between the HPA axis and the brain limbic emotion center, and provide preliminary support that interventions on reducing negative affect may target to reduce the HPA axis and cortisol responses.

5. Limitation

This *meta*-analysis selectively focused on healthy adults, thus the results may not generalize to individuals with major psychiatric or physical illnesses, or to those aged below 18 years or over 65 years. Future reviews could additionally involve patient samples to explore possible relationships between acute stress response and clinical conditions. Due to the paucity of studies on some important physiological and inflammatory biomarkers (e.g., HRV, IL-6), it was not possible to conduct meaningful *meta*-analyses on those measures, which can only be resolved through accumulation of further empirical studies.

6. Conclusion

In summary, we found that salivary cortisol and negative state affects were the most sensitive physiological and psychological markers to acute stress and recovery in the TSST paradigm. SC also received the strongest evidence in terms of association with mood measures during acute stress and recovery. However, existing evidence is lacking for various other potential biomarkers during acute stress. Also, more research is needed to establish the relationship between biological and mood responses during acute stress.

Funding Sources

The project was funded by the Hong Kong Research Grants Council Collaborative Research Fund (C7079-19GF) to T. Lee and the Open Fund from the Guangdong-Hong Kong-Macau Greater Bay Area Centre for Brain Science and Brain-Inspired Intelligence (2019007) to R. Shao.

Author Contributions

I.S.C.M. contributed to paper screening and data extraction, data analysis and interpretation, and manuscript drafting. R.S. contributed to study conceptualization, paper screening and data extraction, data analysis and interpretation, manuscript drafting and revision. T.M.C.L., S-Y.Y, H.W.K., S.X.L, F.Y.L., M.L., Y.K.W. contributed to study conceptualization and manuscript revision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Dr. Clive Wong for advice in R programming. We thank Mr. Marco Se for assistance in articles searching.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yfrne.2022.101050.

References

Aleknaviciute, J., Tulen, J.H., Kamperman, A.M., de Rijke, Y.B., Kooiman, C.G., Kushner, S.A., 2016. Borderline and cluster C personality disorders manifest distinct physiological responses to psychosocial stress. Psychoneuroendocrinology. 72, 131–138.

Allen, A.P., Kennedy, P.J., Cryan, J.F., Dinan, T.G., Clarke, G., 2014. Biological and psychological markers of stress in humans: focus on the Trier Social Stress Test. Neurosci. Biobehav. Rev. 38, 94–124.

Anderson, L.C., Garrett, J.R., Johnson, D.A., Kauffman, D.L., Keller, P.J., Thulin, A., 1984. Influence of circulating catecholamines on protein secretion into rat parotid saliva during parasympathetic stimulation. J. Physiol. 352 (1), 163–171.

Bae, Y.J., Gaudl, A., Jaeger, S., Stadelmann, S., Hiemisch, A., Kiess, W., Willenberg, A., Schaab, M., von Klitzing, K., Thiery, J., Ceglarek, U., 2016 May 1. Immunoassay or LC-MS/MS for the measurement of salivary cortisol in children? Clin. Chem. Lab. Med. (CCLM). 54 (5), 811–822.

Berntson, G.G., Cacioppo, J.T., 2004. Heart rate variability: stress and psychiatric conditions. Dynam. Electrocardiogr. 41 (2), 57–64.

Beurel, E., Nemeroff, C.B., 2014. Interaction of stress, corticotropin-releasing factor, arginine vasopressin and behaviour. Behav. Neurobiol. Stress-related Disorders 67–80.

Böbel TS, Hackl SB, Langgartner D, Jarczok MN, Rohleder N, Rook GA, Lowry CA, Gündel H, Waller C, Reber SO. Less immune activation following social stress in rural vs. urban participants raised with regular or no animal contact, respectively. Proceedings of the National Academy of Sciences. 2018; 115(20): 5259–64.

Bosch, J.A., de Geus, E.J., Veerman, E.C., Hoogstraten, J., Amerongen, A.V., 2003. Innate secretory immunity in response to laboratory stressors that evoke distinct patterns of cardiac autonomic activity. Psychosom. Med. 65 (2), 245–258.

Cărnuță, M., Crișan, L.G., Vulturar, R., Opre, A., Miu, A.C., 2015. Emotional non-acceptance links early life stress and blunted cortisol reactivity to social threat. Psychoneuroendocrinology 51, 176–187.

Carpenter, L.L., Gawuga, C.E., Tyrka, A.R., Lee, J.K., Anderson, G.M., Price, L.H., 2010. Association between plasma IL-6 response to acute stress and early-life adversity in healthy adults. Neuropsychopharmacology 35 (13), 2617–2623.

Chen, I.Y., Jarrin, D.C., Ivers, H., Morin, C.M., 2017. Investigating psychological and physiological responses to the Trier Social Stress Test in young adults with insomnia. Sleep Med. 40, 11–22.

Chong, R.Y., Uhart, M., McCaul, M.E., Johnson, E., Wand, G.S., 2008. Whites have a more robust hypothalamic–pituitary–adrenal axis response to a psychological stressor than blacks. Psychoneuroendocrinology 33 (2), 246–254.
Christian, L.M., Glaser, R., Porter, K., Iams, J.D., 2013. Stress-induced inflammatory responses in women: effects of race and pregnancy. Psychosom. Med. 75 (7), 658.
Coelho, R., Viola, T.W., Walss-Bass, C., Brietzke, E., Grassi-Oliveira, R., 2014. Childhood maltreatment and inflammatory markers: a systematic review. Acta Psychiatr. Scand. 129 (3), 180–192.
Cohen, J., 2013. Statistical power analysis for the behavioral sciences. Academic Press, Routledge.
Dedovic, K., Duchesne, A., Andrews, J., Engert, V., Pruessner, J.C., 2009. The brain and the stress axis: the neural correlates of cortisol regulation in response to stress. Neuroimage 47 (3), 864–871.
Dickerson, S.S., 2008. Emotional and physiological responses to social-evaluative threat. Soc. Pers. Psychol. Compass 2 (3), 1362–1378.
Dickerson, S.S., Gruenewald, T.L., Kemeny, M.E., 2004. When the social self is threatened: Shame, physiology, and health. J. Pers. 72 (6), 1191–1216.
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.
Dutheil, F., de Saint, V.S., Pereira, B., Schmidt, J., Moustafa, F., Charkhabi, M., Bouillon-Minois, J.B., Clinchamps, M., 2021. DHEA as a biomarker of stress: a systematic review and meta-analysis. Frontiers. Psychiatry 12.
Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56 (2), 455–463.
Edelstein, R.S., Yim, I.S., Quas, J.A., 2010. Narcissism predicts heightened cortisol reactivity to a psychosocial stressor in men. J. Res. Pers. 44 (5), 565–572.
Edwards, K.M., Burns, V.E., Ring, C., Carroll, D., 2006. Sex differences in the interleukin-6 response to acute psychological stress. Biol. Psychol. 71 (3), 236–239.
Egger, M., Smith, G.D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. Br. Med. J. 315 (7109), 629–634.
Fanselow, M.S., Dong, H.W., 2010. Are the dorsal and ventral hippocampus functionally distinct structures? Neuron 65 (1), 7–19.
Fehr, F.S., Stern, J.A., 1970. Peripheral physiological variables and emotion: the James-Lange theory revisited. Psychol. Bull. 74 (6), 411–424.
Ferreira, N.D., Gehin, C., Massot, B., 2021 Feb 1. A review of methods for non-invasive heart rate measurement on wrist. IRBM. 42 (1), 4–18.
Flandreau, E.I., Ressler, K.J., Owens, M.J., Nemeroff, C.B., 2012. Chronic overexpression of corticotropin-releasing factor from the central amygdala produces HPA axis hyperactivity and behavioral anxiety associated with gene-expression changes in the hippocampus and paraventricular nucleus of the hypothalamus. Psychoneuroendocrinology 37 (1), 27–38.
Fomicheva, E.E., Nemirovich-Danchenko, E.A., Korneva, E.A., 2004. Immunoprotective effects of prolactin during stress-induced immune dysfunction. Bull. Exp. Biol. Med. 137 (6), 544–547.
Fulford, A.J., Harbuz, M.S., 2005. An introduction to the HPA axis. In: Steckler, T., Kalin, N.H., Reul, J.M.H.M. (Eds.), Handbook of Stress and The Brain - Part 1: The Neurobiology of Stress. Elsevier, pp. 43–65.
Gallagher, P., Leitch, M.M., Massey, A.E., McAllister-Williams, R.H., Young, A.H., 2006 Sep. Assessing cortisol and dehydroepiandrosterone (DHEA) in saliva: effects of collection method. J. Psychopharmacol. 20 (5), 643–649.
Garcia-Rubio, M.J., Espin, L., Hidalgo, V., Salvador, A., Gomez-Amor, J., 2017. Autonomic markers associated with generalized social phobia symptoms: heart rate variability and salivary alpha-amylase. Stress 20 (1), 61–68.
Gasperin, D., Netuveli, G., Dias-da-Costa, J.S., Pattussi, M.P., 2009. Effect of psychological stress on blood pressure increase: a meta-analysis of cohort studies. Cad. Saude Publica. 25 (4), 715–726.
Golatowski, C., Salazar, M.G., Dhople, V.M., Hammer, E., Kocher, T., Jehmlich, N., Völker, U., 2013 Apr. Comparative evaluation of saliva collection methods for proteome analysis. Clin. Chim. Acta 18 (419), 42–46.
Goldberger, J.J., 1999. Sympathovagal balance: how should we measure it? Am. J. Physiol.-Heart Circulatory Physiol. 276 (4), H1273–H1280.
Goodman, W., Janson, J., Wolf, J., 2017. Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test. Psychoneuroendocrinology 80, 26–35.
Goodwin, R., Kaniasty, K., Sun, S., Ben-Ezra, M., 2017. Psychological distress and prejudice following terror attacks in France. J. Psychiatr. Res. 91, 111–115.
Harrewijn, A., Vidal-Ribas, P., Clore-Gronenborn, K., Jackson, S.M., Pisano, S., Pine, D.S., Stringaris, A., 2020. Associations between brain activity and endogenous and exogenous cortisol–A systematic review. Psychoneuroendocrinology 120, 104775.
Helminen, E.C., Morton, M.L., Wang, Q., Felver, J.C., 2019. A meta-analysis of cortisol reactivity to the Trier Social Stress Test in virtual environments. Psychoneuroendocrinology 110, 104437.
Hendrawan, D., Yamakawa, K., Kimura, M., Murakami, H., Ohira, H., 2012. Executive functioning performance predicts subjective and physiological acute stress reactivity: preliminary results. Int. J. Psychophysiol. 84 (3), 277–283.
Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane handbook for systematic reviews of interventions. 2nd edition. Chichester: John Wiley & Sons. 2019.
Higgins, J.P., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. Stat. Med. 21 (11), 1539–1558.
Huedo-Medina, T.B., Sánchez-Meca, J., Marin-Martinez, F., Botella, J., 2006. Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? Psychol. Methods 11 (2), 193.
IntHout, J., Ioannidis, J., Borm, G.F., 2014. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med. Res. Method. 14 (1), 1–2.
Izawa, S., Sugaya, N., Shirotsuki, K., Yamada, K.C., Ogawa, N., Ouchi, Y., Nagano, Y., Suzuki, K., Nomura, S., 2008. Salivary dehydroepiandrosterone secretion in response to acute psychosocial stress and its correlations with biological and psychological changes. Biol. Psychol. 79 (3), 294–298.
Izawa, S., Sugaya, N., Kimura, K., Ogawa, N., Yamada, K.C., Shirotsuki, K., Mikami, I., Hirata, K., Nagano, Y., Nomura, S., 2013. An increase in salivary interleukin-6 level following acute psychosocial stress and its biological correlates in healthy young adults. Biol. Psychol. 94 (2), 249–254.
Jackson, D., Law, M., Rücker, G., Schwarzer, G., 2017. The Hartung-Knapp modification for random-effects meta-analysis: a useful refinement but are there any residual concerns? Stat. Med. 36 (25), 3923–3934.
Janusek, L.W., Tell, D., Gaylord-Harden, N., Mathews, H.L., 2017. Relationship of childhood adversity and neighborhood violence to a proinflammatory phenotype in emerging adult African American men: an epigenetic link. Brain Behav. Immun. 60, 126–135.
Joyner, M.J., Charkoudian, N., Wallin, B.G., 2008. A sympathetic view of the sympathetic nervous system and human blood pressure regulation. Exp. Physiol. 93 (6), 715–724.
Juster, R.P., McEwen, B.S., Lupien, S.J., 2010. Allostatic load biomarkers of chronic stress and impact on health and cognition. Neurosci. Biobehav. Rev. 35 (1), 2–16.
Kennedy, P.J., Cryan, J.F., Quigley, E.M., Dinan, T.G., Clarke, G., 2014. A sustained hypothalamic–pituitary–adrenal axis response to acute psychosocial stress in irritable bowel syndrome. Psychol. Med. 44 (14), 3123–3134.
Kern, S., Oakes, T.R., Stone, C.K., McAuliff, E.M., Kirschbaum, C., Davidson, R.J., 2008. Glucose metabolic changes in the prefrontal cortex are associated with HPA axis response to a psychosocial stressor. Psychoneuroendocrinology 33 (4), 517–529.
Khera, A., McGuire, D.K., Murphy, S.A., Stanek, H.G., Das, S.R., Vongpatanasin, W., Wians, F.H., Grundy, S.M., de Lemos, J.A., 2005. Race and gender differences in C-reactive protein levels. J. Am. Coll. Cardiol. 46 (3), 464–469.
Kimura, K., Izawa, S., Sugaya, N., Ogawa, N., Yamada, K.C., Shirotsuki, K., Mikami, I., Hirata, K., Nagano, Y., Hasegawa, T., 2013. The biological effects of acute psychosocial stress on delay discounting. Psychoneuroendocrinology 38 (10), 2300–2308.
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.
Kische, H., Ollmann, T.M., Voss, C., Hoyer, J., Rückert, F., Pieper, L., Kirschbaum, C., Beesdo-Baum, K., 2021. Associations of saliva cortisol and hair cortisol with generalized anxiety, social anxiety, and major depressive disorder: an epidemiological cohort study in adolescents and young adults. Psychoneuroendocrinology 126, 105167.
Klumbies, E., Braeuer, D., Hoyer, J., Kirschbaum, C., 2014. The reaction to social stress in social phobia: discordance between physiological and subjective parameters. PLoS One 9 (8), e105670.
Koray M, Dülger O, Ak G, Horasanli S, Ü çok A, Tanyeri H, Badur S. The evaluation of anxiety and salivary cortisol levels in patients with oral lichen planus. Oral Diseases. 2003;9(6):298-301.
Lackschewitz, H., Hüther, G., Kröner-Herwig, B., 2008. Physiological and psychological stress responses in adults with attention-deficit/hyperactivity disorder (ADHD). Psychoneuroendocrinology 33 (5), 612–624.
Lakens, D., 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front. Psychol. 4, 863.
Lee, T.M., Leung, M.K., Hou, W.K., Tang, J.C., Yin, J., So, K.F., Lee, C.F., Chan, C.C., 2012. Distinct neural activity associated with focused-attention meditation and loving-kindness meditation. PLoS One 7 (8), e40054.
Lemaire, J.J., Frew, A.J., McArthur, D., Gorgulho, A.A., Alger, J.R., Salomon, N., Chen, C., Behnke, E.J., De Salles, A.A., 2011. White matter connectivity of human hypothalamus. Brain Res. 1371, 43–64.
Lennartsson, A.K., Jonsdottir, I.H., 2011. Prolactin in response to acute psychosocial stress in healthy men and women. Psychoneuroendocrinology 36 (10), 1530–1539.
Lennartsson, A.K., Kushnir, M.M., Bergquist, J., Jonsdottir, I.H., 2012. DHEA and DHEA-S response to acute psychosocial stress in healthy men and women. Biol. Psychol. 90 (2), 143–149.
Liu, J.J., Ein, N., Peck, K., Huang, V., Pruessner, J.C., Vickers, K., 2017 Aug. Sex differences in salivary cortisol reactivity to the Trier Social Stress Test (TSST): a meta-analysis. Psychoneuroendocrinology 1 (82), 26–37.
McEwen, B.S., 1998. Stress, adaptation, and disease: allostasis and allostatic load. Ann. N. Y. Acad. Sci. 840 (1), 33–44.
Merswolken, M., Deter, H.C., Siebenhuener, S., Orth-Gomér, K., Weber, C.S., 2013. Anxiety as predictor of the cortisol awakening response in patients with coronary heart disease. Int. J. Behav. Med. 20 (3), 461–467.
Merz, C.J., Wolf, O.T., 2015. Examination of cortisol and state anxiety at an academic setting with and without oral presentation. Stress 18 (1), 138–142.
Mikolajczak, M., Roy, E., Luminet, O., De Timary, P., 2008. Resilience and hypothalamic-pituitary-adrenal axis reactivity under acute stress in young men. Stress 11 (6), 477–482.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. 151 (4), 264–269.
Monteleone, A.M., Patriciello, G., Ruzzi, V., Cimino, M., Del Giorno, C., Steardo Jr, L., Monteleone, P., Maj, M., 2018 Sep. Deranged emotional and cortisol responses to a psychosocial stressor in anorexia nervosa women with childhood trauma exposure: evidence for a "maltreated ecophenotype"? J. Psychiatr. Res. 1 (104), 39–45.
Munro, C.A., Oswald, L.M., Weerts, E.M., McCaul, M.E., Wand, G.S., 2005. Hormone responses to social stress in abstinent alcohol-dependent subjects and social drinkers with no history of alcohol dependence. Alcohol. Clin. Exp. Res. 29 (7), 1133–1138.

Nater, U.M., Rohleder, N., 2009. Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research. Psychoneuroendocrinology 34 (4), 486–496.
Nolen-Hoeksema S, Aldao A. Gender and age differences in emotion regulation strategies and their relationship to depressive symptoms. Personality and Individual Differences. 201;51(6):704-8.
Noto, Y., Sato, T., Kudo, M., Kurata, K., Hirota, K., 2005. The relationship between salivary biomarkers and state-trait anxiety inventory score under mental arithmetic stress: a pilot study. Anesth. Analg. 101 (6), 1873–1876.
Otte, C., Hart, S., Neylan, T.C., Marmar, C.R., Yaffe, K., Mohr, D.C., 2005. A meta-analysis of cortisol response to challenge in human aging: importance of gender. Psychoneuroendocrinology 30 (1), 80–91.
Petrakova, L., Doering, B.K., Vits, S., Engler, H., Rief, W., Schedlowski, M., Grigoleit, J.S., 2015. Psychosocial stress increases salivary alpha-amylase activity independently from plasma noradrenaline levels. PLoS One 10 (8), e0134561.
Pizzagalli, D.A., 2014. Depression, stress, and anhedonia: toward a synthesis and integrated model. Annu. Rev. Clin. Psychol. 10, 393–423.
Purves, D., Augustine, G.J., Fitzpatrick, D., Katz, L., LaMantia, A.S., McNamara, J.O., Williams, S.M., 2001. Neuroscience, 2nd edition. Sinauer Associates, Sunderland (MA).
Raff, H., Homar, P.J., Burns, E.A., 2002 Jan 1. Comparison of two methods for measuring salivary cortisol. Clin. Chem. 48 (1), 207–208.
Rimmele, U., Zellweger, B.C., Marti, B., Seiler, R., Mohiyeddini, C., Ehlert, U., Heinrichs, M., 2007. Trained men show lower cortisol, heart rate and psychological responses to psychosocial stress compared with untrained men. Psychoneuroendocrinology 32 (6), 627–635.
Schneiderman, N., Ironson, G., Siegel, S.D., 2005. Stress and health: psychological, behavioral, and biological determinants. Annu. Rev. Clin. Psychol. 1, 607–628.
Schwarzer G, Carpenter JR, Rücker G. Small-study effects in meta-analysis. Meta-analysis with R. Switzerland: Springer Cham; 2015.
Seddon, J., Rodriguez, V., 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.
Seeman, T.E., McEwen, B.S., Rowe, J.W., Singer, B.H., 2001. Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. Proc. Natl. Acad. Sci. 98 (8), 4770–4775.
Sequeira, H., Hot, P., Silvert, L., Delplanque, S., 2009. Electrical autonomic correlates of emotion. Int. J. Psychophysiol. 71 (1), 50–56.
Shacham, S., 1983. A shortened version of the Profile of Mood States. J. Pers. Assess. 47 (3), 305–306.
Shirtcliff, E.A., Granger, D.A., Schwartz, E., Curran, M.J., 2001 Feb 1. Use of salivary biomarkers in biobehavioral research: cotton-based sample collection methods can interfere with salivary immunoassay results. Psychoneuroendocrinology 26 (2), 165–173.
Slavich, G.M., Irwin, M.R., 2014. From stress to inflammation and major depressive disorder: a social signal transduction theory of depression. Psychol. Bull. 140 (3), 774.
Sloman, L., Gilbert, P., Hasey, G., 2003. Evolved mechanisms in depression: the role and interaction of attachment and social rank in depression. J. Affect. Disord. 74 (2), 107–121.
Spiga, F., Lightman, S.L., 2015. Dynamics of adrenal glucocorticoid steroidogenesis in health and disease. Mol. Cell. Endocrinol. 408, 227–234.
Steptoe, A., Hamer, M., Chida, Y., 2007. The effects of acute psychological stress on circulating inflammatory factors in humans: a review and meta-analysis. Brain Behav. Immun. 21 (7), 901–912.
Sterne, J.A., Sutton, A.J., Ioannidis, J.P., Terrin, N., Jones, D.R., Lau, J., Carpenter, J., Rücker, G., Harbord, R.M., Schmid, C.H., Tetzlaff, J., 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Br. Med. J. 343.
Sterne, J.A., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., Viswanathan, M., Henry, D., Altman, D.G., Ansari, M.T., Boutron, I., Carpenter, J.R., 2016. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. Br. Med. J. 355.
Storm, H., Myre, K., Rostrup, M., Stokland, O., Lien, M.D., Raeder, J.C., 2002. Skin conductance correlates with perioperative stress. Acta Anaesthesiol. Scand. 46 (7), 887–895.
Sugaya, N., Izawa, S., Kimura, K., Ogawa, N., Yamada, K.C., Shirotsuki, K., Mikami, I., Hirata, K., Nagano, Y., Nomura, S., Shimada, H., 2012. Adrenal hormone response and psychophysiological correlates under psychosocial stress in individuals with irritable bowel syndrome. Int. J. Psychophysiol. 84 (1), 39–44.
Szabo, Y.Z., Slavish, D.C., Graham-Engeland, J.E., 2020. The effect of acute stress on salivary markers of inflammation: a systematic review and meta-analysis. Brain Behav. Immun. 88, 887–900.
Takai, N., Yamaguchi, M., Aragaki, T., Eto, K., Uchihashi, K., Nishikawa, Y., 2004. Effect of psychological stress on the salivary cortisol and amylase levels in healthy young adults. Arch. Oral Biol. 49 (12), 963–968.
Uhart, M., Oswald, L., McCaul, M.E., Chong, R., Wand, G.S., 2006. Hormonal responses to psychological stress and family history of alcoholism. Neuropsychopharmacology 31 (10), 2255–2263.
van den Ouweland, J.M., Kema, I.P., 2012. The role of liquid chromatography–tandem mass spectrometry in the clinical laboratory. J. Chromatogr. B 883, 18–32.
Viechtbauer, W., 2005. Bias and efficiency of meta-analytic variance estimators in the random-effects model. J. Educ. Behav. Stat. 30 (3), 261–293.
Viechtbauer, W., Cheung, M.W., 2010. Outlier and influence diagnostics for meta-analysis. Res. Synth. Methods 1 (2), 112–125.
Watson, D., Clark, L., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. J. Pers. Soc. Psychol. 54 (6), 1063–1070.
Wearne, T.A., Lucien, A., Trimmer, E.M., Logan, J.A., Rushby, J., Wilson, E., Filipčíková, M., McDonald, S., 2019. Anxiety sensitivity moderates the subjective experience but not the physiological response to psychosocial stress. Int. J. Psychophysiol. 141, 76–83.
Wemm, S.E., Wulfert, E., 2017. Effects of acute stress on decision making. Appl. Psychophysiol. Biofeedback 42 (1), 1–2.
Whittle, S., Yücel, M., Yap, M.B., Allen, N.B., 2011. Sex differences in the neural correlates of emotion: evidence from neuroimaging. Biol. Psychol. 87 (3), 319–333.
Yamakawa, K., Matsunaga, M., Isowa, T., Ohira, H., 2015. Serotonin transporter gene polymorphism modulates inflammatory cytokine responses during acute stress. Sci. Rep. 5 (1), 1–9.
Zhang, H., Yao, Z., Lin, L., Sun, X., Shi, X., Zhang, L., 2019. Early life stress predicts cortisol response to psychosocial stress in healthy young adults. PsyCh Journal. 8 (3), 353–362.