Experimental stress induction in children and adolescents with the Trier Social Stress Test (TSST): a systematic review and meta-analysis

Saskia Seel[1,3], Bernhard Pastötter[2,3], & Gregor Domes[1,3]

[1] Biological and Clinical Psychology, Department of Psychology, University of Trier, Trier, Germany

[2] Cognitive Psychology, Department of Psychology, University of Trier, Trier, Germany

[3] Institute for Cognitive and Affective Neuroscience, University of Trier, Trier, Germany

Corresponding author: Gregor Domes, University of Trier, Department of Biological and Clinical Psychology, Johanniterufer 15, 54290 Trier, Germany. *E-mail address*: domes@uni-trier.de; *Phone*: +49 651 201 2929; *Fax*: +49 651 201 2934

Abstract

The Trier Social Stress Test (TSST) is one of the most frequently used laboratory stressors. Previous meta-analyses concluded that the TSST induces moderate but robust psychobiological stress responses in both adult and, recently, youth samples. With the present study, we focused on the salivary cortisol response and performed both a systematic review comprising 143 studies and a meta-analysis with 91 studies (N=8291 participants) (a) to investigate which and how different TSST variants have been used in children and adolescents (b) to provide an updated overview of the reported effect sizes, and (c) to identify previously reported and new potential moderators. Our results showed that TSST characteristics have been reported inconsistently, with generally similar administration characteristics for the major TSST versions TSST for Children (TSST-C), TSST Modified (TSST-M), and the original TSST regarding participant exclusion, administration time, preparation, total TSST duration, and jury set-up. After correcting for publication bias using PEESE residuals, we found the TSST to be an effective tool for inducing a salivary cortisol stress response (*Hedge's g* = 0.56, p<.0001). Moderation analyses revealed that participant gender, number of judges, and the TSST's total duration were significant moderators. Exploratively speaking, higher baseline cortisol levels predicted lower baseline-peak reactivity. Overall, the TSST (with its specific adaptations) is a suitable instrument for inducing acute psychobiological stress responses in children and adolescents. In the interests of open and reproducible science, establishing reporting standards for both methods and results would be desirable for future studies to enable more robust moderation analyses.

Keywords.

Trier Social Stress Test; TSST; children; adolescents; youth; stress induction; cortisol; HPA-axis

1. Introduction

The basic idea of the Trier Social Stress Test (TSST; Kirschbaum et al., 1993) as a laboratory procedure is to impose social-evaluative threat on an individual by generating anticipation and the perception of both uncontrollability and negative evaluation from others. This is supposed to lead to psychological and physiological stress responses in these individuals (Dickerson and Kemeny, 2004). The original design was developed for an adult sample (Kirschbaum et al., 1993) and consists of a preparation phase for a previously announced public speaking task, the public speaking task itself, and a surprise verbal arithmetic task. A panel of "judges" presides over all these phases and provides elements of social-evaluative threat.

The TSST is known as the gold standard for laboratory stress induction (Allen et al., 2017). Regarding the psychological stress reaction, the TSST has been shown to increase self-reported stress, anxiety, and negative mood in adult participants (Allen et al., 2014; von Dawans et al., 2011). In terms of physiological stress reactions, stress triggers the activation of the autonomic nervous system (ANS) and hypothalamic-pituitary-adrenal (HPA) axis. In the first "faster" route, a "fight or flight response" is set off by the secretion of adrenaline and noradrenaline via the sympathomedullary pathway (Allen et al., 2014; Labuschagne et al., 2019). Aside from the salivary alpha-amylase concentration (Gordis et al., 2006), the TSST has been shown to cause a significant increase in other relevant ANS activation markers such as, but not limited to, heart rate (Childs et al., 2006; Kudielka et al., 2004) and respiratory sinus arrythmia (Ho et al., 2020; Rozenman et al., 2018). In the second "slower" route, the hypothalamus is stimulated to secrete corticotrophin-releasing hormone (CRH), which in turn promotes the release of adrenocorticotrophic hormone (ACTH) from the pituitary gland, through which the secretion of cortisol from the adrenal cortex is set off (Allen et al., 2014; Kudielka et al., 2007). Salivary cortisol is a commonly used (Goodman et al., 2017) and reliable (Hellhammer et al., 2009) marker for HPA axis activation, and the TSST has been repeatedly shown to induce significant increases in adults (Kirschbaum et al., 1993; von Dawans et al., 2011), adolescents (Cameron et al., 2017; Gunnar et al., 2021; Kelly et al., 2007; Yim et al., 2015), and children (Buske-Kirschbaum et al., 1997; Yim et al., 2015). In their meta-analysis of acute stress effects in healthy adults, Dickerson and Kemeny (2004) estimated stressors like the TSST that combined public speaking and cognitive tasks to have a "larger average effect size than the other stressor categories, […] with an effect size over two times as large as any other types of stressor" (p. 369). A more recent meta-analysis by Gu et al. (2022), which examined acute stress effects in TSST studies on healthy adults, reported similar results.

In the 30 years since its original development, the TSST has been adapted for application in different age groups (typically adults, teenagers, or children) and experimental settings (solo or group), resulting in several specific TSST versions used by different research groups. These versions differ primarily in the topic of the public speech task, the durations of the different phases, the number of judges, and the story context in which the judges are introduced in (see Table 1). The versions, incidentally, often undergo additional adaptation to fulfill the needs of a specific

research group, potentially compromising its validity and reproducibility (Labuschagne et al., 2019). In their meta-analysis, Narvaez Linares et al. (2020) also highlighted significant differences across studies with adult samples and the need for a standardized reporting scheme to transparently communicate the TSST protocol used in individual studies. In their meta-analysis of 186 TSST studies on healthy adults, Goodman et al. (2017) analyzed the moderating effects of common variations of the classic TSST protocol. They identified several significant moderators influencing participants' cortisol reactivity peaking around 35-45 minutes after stressor onset. Notable factors included the administration time of the TSST protocol (larger effects when administered between 2pm and 5pm), the acclimation time for the participant before the start of the protocol (larger effects with acclimation time between 16 and 30 minutes), the nature of tasks preceding the TSST protocol (larger effects in studies with stressful tasks beforehand), the magnitude of the subtraction number in the arithmetic task (larger effects in studies using larger subtraction numbers), as well as the composition of the judging panel in terms of number, gender, and facial feedback (larger effects in conjunction with mixed-gender juries consisting of two judges with neutral facial feedback).

Several years after the publication of the adult version of the TSST, the first adaptation for children was used to induce stress and investigate the acute response of the HPA axis in pediatric study samples (Buske-Kirschbaum et al., 1997). Subsequently, like in adult research, validated TSST protocols for children and adolescents have undergone various adaptations. For example, singing in front of the judges replaced the classic public speaking task (Brand et al., 2011), performance pressure was heightened by introducing the possibility of receiving an unwanted gift due to an unfavorable judgement by the jury (de Veld et al., 2014a, 2014b), and when the jury was seated behind a one-way mirror (Hostinar et al., 2015a, 2015b).

In their recent meta-analysis, which focused on studies using the TSST with participants aged seven to 17 years, Seddon et al. (2020) considered all modifications of the original protocol such as these to provide a broad overview of the field. They assessed several biomarkers such as salivary cortisol, salivary alpha-amylase, heart rate, systolic and diastolic blood pressure, cardiac output, HRV, pre-ejection period, and respiratory sinus arrhythmia, and tested moderators of the stress response. Including data from 57 publications up to February 2018, Seddon et al. (2020) confirmed the TSST as an effective stressor for a youth sample with a moderate overall effect on cortisol levels after accounting for publication bias (d'=.47). The researchers identified age (younger children revealing a greater cortisol increase pre-post TSST than adolescents), sample type (community samples showing stronger increases than clinical samples), gender of the judges (mixed-gendered juries eliciting greater cortisol increases), total TSST duration (longer TSST duration eliciting higher increases), and time of administration (TSST in the afternoon) as significant moderators of cortisol reactivity.

Table 1. Overview of commonly used versions of the TSST concept.

			Phase durations			Task in detail		Judges		
Version	Age range	Cover story judges	Prep.	Speech	Arith.	Speech	Arithmetic	Number	Gender	Behavior
Trier Social Stress Test (TSST; Kirschbaum et al., 1993)	Adults (min. 18)	Selection committee, trained to monitor non-verbal behavior, video, and voice analysis announced	10	5	5	Personal interview with a company's selection committee as a job applicant for a vacant position	Everyone: 1022 – 13	3	Mixed	Neutral
Trier Social Stress Test for Children (TSST-C; Buske-Kirschbaum et al., 1997)	Children (9 – 14)	Evaluative audience	5	5	5	Complete a story from a given story stem with a "good" narrative, graded performance	9–11-year-olds: 758 – 7 12-14-year-olds: 1023 – 13	2	Not specified	Friendly, supportive
Trier Social Stress Test in Virtual Reality (TSST-VR; Kelly et al., 2007)	Teenagers (min. 15), Adults	Evaluative audience	5	5	5	Speech about themselves to impress employers	Everyone: 2013 – 17	5	Mixed	First neutral, progressively bored
Trier Social Stress Test modified (TSST-M; Yim et al., 2010)	Children (9 - 12), Adults (18 - 25)	Evaluative audience, video and voice analysis announced	3	6	4	Introduction to an imagined new class of peers, saying good and bad things about themselves	9-12-year-olds: 1027 – 5 18-25-year-olds: 1027 – 13	2	Mixed	Neutral
Trier Social Stress Test for Groups (TSST-G; von Dawans et al., 2011)	Adults (min. 18)	Selection committee, trained to monitor non-verbal behavior, video, and voice analysis announced	10	2 Group: 12	1.2 Group: 8	Group interview with a company's selection committee as a job applicant for a vacant position	Everyone: unique number – 16	2	Mixed	Neutral
Trier Social Stress Test for Teenagers (TSST-T; Cameron et al., 2017)	Teenagers (min. 12)	Evaluative audience, video and voice analysis announced	5	5	3	Complete a story from a given story stem with a "good" narrative, graded performance	Everyone: 2037 – 13	2	Female	Still faced
Trier Social Stress Test Online (TSST-OL; Gunnar et al., 2021)	Teenagers (15 – 16)	Evaluative audience	5	5	5	Introduction to a class of peers, saying good and bad things about themselves	Everyone: 938 – 13	2	Mixed	Neutral

Note. Age range in years. Phase durations (preparation, public speaking task, arithmetic task) in minutes. Arithmetic task is always a serial subtraction task.

With the present work we want to update and expand upon the findings of Seddon et al. (2020) focusing on the stress related salivary cortisol response. In addition, a comprehensive overview of the application and methodology of the TSST in published studies with children and adolescents will be provided, similar to the work of Narvaez Linares et al. (2020) in the context of TSSTs with adult samples. To achieve this, we conducted a systematic review on studies using variations of the TSST procedure with healthy children and adolescents aged between 6 and 18 years of age. We evaluated these in terms of reported methodological characteristics to identify specific protocol variations that may influence the observed responses. Furthermore, we conducted a meta-analysis to estimate the overall effect of the TSST on the salivary cortisol (baseline-to-peak) responses in children and adolescents. This analysis also included analyses of established moderators reported previously (see Goodman et al., 2017; Seddon et al., 2020) and potentially new moderators linked to various TSST versions and experimenter variables for the present non-adult samples.

2. Methods

2.1. Search strategy and study selection

On September 26th 2023, we screened MEDLINE (PubMed), PsychINFO (Ovid), and Web of Science (Core collection) using the search term [("TSST" OR "Trier Social Stress") AND (children OR adolescents OR youth) NOT “toxic shock”] to capture published literature specifically related to the TSST in children and adolescents. References were exported to and managed in Zotero (www.zotero.org), which we used to identify and remove duplicates. Following the inclusion/exclusion criteria below, studies were first screened based on the titles and abstracts. Using the same criteria, full-text articles were then assessed for eligibility (see PRISMA flow chart in Fig. 1).

2.2. Inclusion and exclusion criteria

Our meta-analysis employed a two-stepped inclusion/exclusion process. Firstly, peer-reviewed studies were included in the systematic review if (1) the authors had carried out a TSST procedure or applied slight modifications still possessing the elements of evaluative threat (panel vs. video) and uncontrollability (arithmetic task after speech task), (2) the samples in the studies were aged between 6 and 18 years of age, and (3) the authors reported salivary cortisol as a measure for the physiological stress response. Studies were excluded if they (1) were not published in English, (2) only featured a clinical sample, or (3) included additional stressors that could influence the salivary cortisol output triggered by the TSST.

Secondly, data from this pool was included in our meta-analysis if it was derived from studies that (1) used a TSST setup meant to produce stress (in comparison to a control setup or a setup with reduced stress), (2) observed a healthy[1] (control) sample with information about the sample size, and (3) reported mean cortisol values with standard deviation or standard error for at least one baseline sampling time and one post TSST sampling time. Data was excluded if (1) cortisol data was reported as aggregated measures only (AUCi or AUCg) or (2) relevant within-group variances (SD or SEM) were unavailable[2].

2.3. Data extraction, coding, and preparation

The first author performed all data extraction and moderator coding. The first 100 papers were used as training material to settle uncertainties in paper classification and data coding through discussion with the last author.

All titles and abstracts were screened for studies using the TSST with a youth sample. For the systematic review, full articles were coded for the participants' age range , the time window of the individual assessments, the adaptation time window in the lab before the start of the TSST routine, the saliva sampling method (i.e., Salivettes), the total number of saliva samples taken during the assessment, the saliva sampling times relative to the start of the TSST, the TSST version used (i.e., TSST, TSST-C, TSST-M), the duration of the phases of the TSST used, details about the judges (number of judges, gender composition, behavior toward the participants), details about the experimenter (gender, behavior toward the participants, double role as judge), details about presented data (reported responder rates to the TSST, the responder rate itself, the responder criterium used), and exclusion criteria used (in reference to the participants themselves and, if applicable, their parents).

For the meta-analysis, usable studies from the qualitative data pool were screened for healthy control sample data and coded for descriptive data (sample size, gender composition, mean age) of the healthy control sample, as well as all available salivary cortisol data (mean and error by sampling time, unit used). Data provided in graphs was extracted using the WebPlotDigitizer (Rohatgi, 2022), while log-transformed data was transformed into approximal raw data using methods described by Higgins and Green (2008). If the data provided (reported via table or graph) was not usable for this meta-analysis or deemed invalid after transformation (eg. too high for the essay used), the first author contacted the corresponding author for the required raw data. If the data was provided by mid-October 2023, it was included in the quantitative analysis.

[1] All samples were comprised of participants without a clinical diagnosis, although some subsamples were somewhat burdened by research-specific factors. We always chose the least burdened and closest to healthy sub-sample.
[2] For one study, we could only extract baseline cortisol and cortisol reactivity data. We used formulas provided by Higgins and Green (2008) to calculate the cortisol peak value.

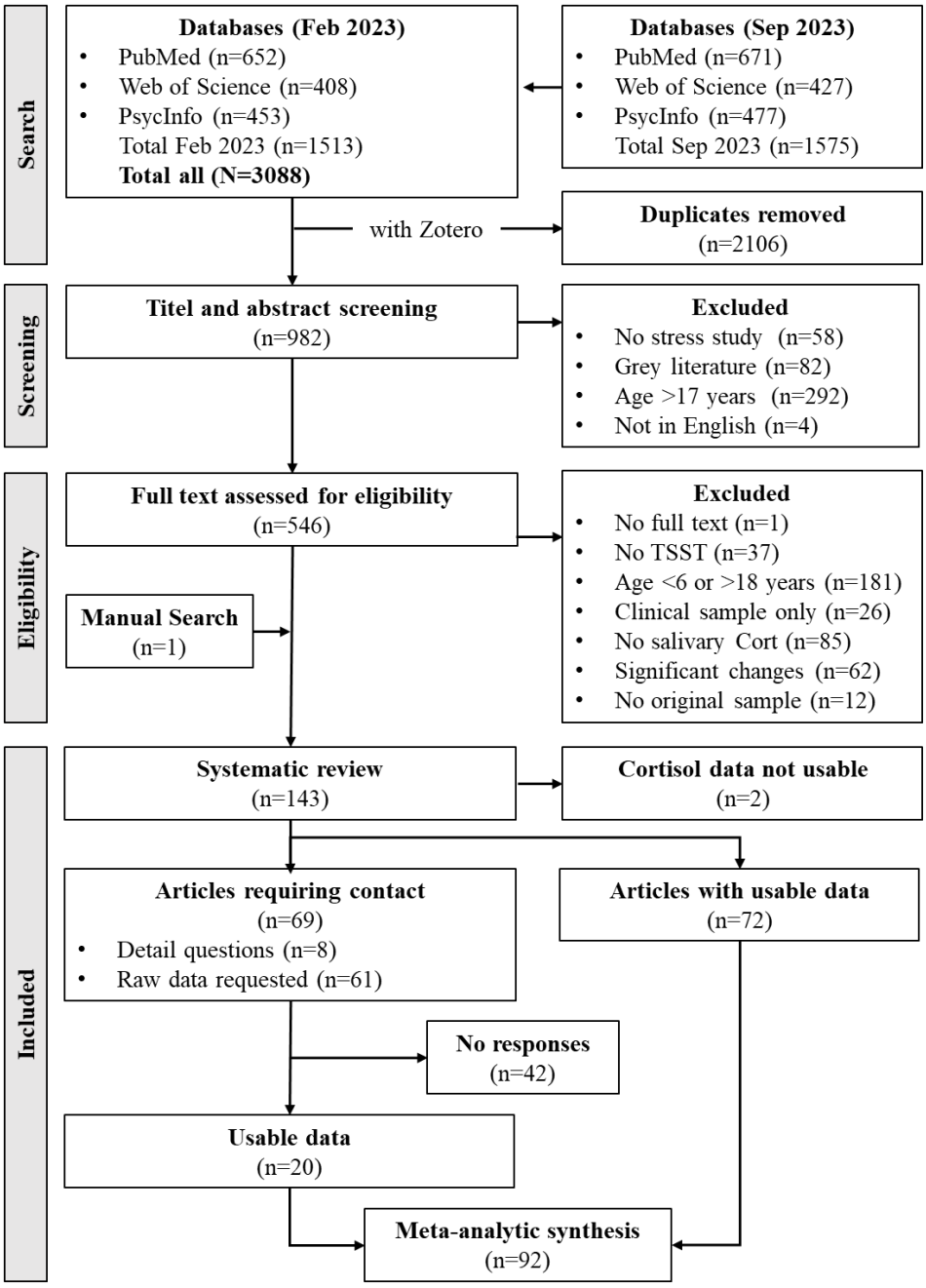


Figure 1. PRISMA flow chart on study selection.

2.4. Data analysis

2.4.1 Calculation of effect sizes

The goal of this meta-analysis was to assess baseline to peak differences in salivary cortisol before and after participation in the TSST. A standardized mean difference (SMD or Cohen's *d*)

effect size was therefore used. The standardization of mean differences utilized the pooled standard deviation of baseline and peak cortisol responses. To calculate the standard errors of SMDs, the correlation between baseline and peak cortisol responses was set to medium r=.3 (see Borenstein et al., 2011). To mitigate any small sample bias in effects, all Cohen's *d* were transformed into Hedge's *g*, and the standard error of *g* were calculated. A Hedge's *g* higher than zero indicates that salivary cortisol levels were higher after TSST exposure in comparison to the baseline, and values lower than zero indicate that salivary cortisol levels fell after the TSST.

To facilitate calculation, we defined the standard measuring unit for salivary cortisol in this meta-analysis as nmol/l. Raw data provided in other measuring units (i.e., ng/ml or µg/dl) were transformed into nmol/l. Provided standard errors were transformed into standard deviations.

2.4.2 Heterogeneity

In the meta-analysis, random-effects modeling was employed to account for potential heterogeneity among the included studies. The Q statistic was calculated to assess the presence of heterogeneity among the effect sizes. The Q statistic is a measure of the dispersion of effect sizes and is calculated as the weighted sum of squared differences between individual study effects and the overall effect. A significant Q statistic indicates the presence of variability beyond what would be expected by chance alone, prompting further exploration of potential sources of heterogeneity through subgroup analyses or meta-regression. Parameters in random-effects modeling were estimated using a restricted (residual) maximum likelihood (REML) estimation model. For all analyses, a cutoff of $p < .05$ was used for significance. Our analyses were conducted in R (R Core Team, 2021) using metafor (Viechtbauer, 2010), robumeta (Fisher et al., 2023), weightr (Coburn & Vevea, 2019), boot (Canty & Ripley, 2022), forcats (Wickham, 2023), clubSandwich (Pustejovsky, 2023), e1071 (Meyer et al., 2023), matrixStats (Bengtsson, 2023), psych (Revelle, 2023), MASS (Venables & Ripley, 2002), rstatix (Kassambara, 2023b), esc (Lüdecke, 2019), ggpubr (Kassambara, 2023a), and pwr (Champely, 2020) packages.

2.4.3 Moderators

The following variables were included in our analyses as potential moderators: participant age (both as a continuous variable taken from the mean age of the included sample(s), and categorical depending on the mean age and standard deviation of the included study sample, with samples classed as "children" with a sum up to 12 and samples classed "adolescents" with a differential over 12), gender of the included sample (male, female, or mixed genders), gender ration of the judges (all female presenting, all male presenting, mixed, same as participant, opposite to participant), number of judges (as a categorical variable), and the TSST's total duration (as a continuous variable) as seen in Seddon et al. (2020) in attempt to replicate their findings. Additionally, the TSST version used (as a categorical variable), the judges' facial feedback (positive/encouraging, neutral, negative, progressively negative), the acclimation time to the

laboratory setting before initiating the TSST (as a continuous variable), the experimenter's gender (always female presenting, always male presenting, mixed gender presentation, same as participant, opposite to participant), and their interaction behavior towards the participant (friendly, neutral, negative) were included to extend upon findings reported in Goodman et al. (2017). As advised by Gartlehner et al. (2011), moderator analyses were conducted if there were at least four effect sizes per moderator level.

2.4.4 Publication bias

Publication bias was tested to assess the possibility of biased results due to systemically missing studies (e.g., small study effect) in the selected studies. Firstly, a meta-regression using the year of publication as a moderator was employed to assess time-based publication bias. Secondly, we ran Egger's regression asymmetry test. Thirdly, precision-effect test (PET) and precision-effect estimate with standard errors (PEESE) were utilized. In case PET-PEESE indicated significant publication bias, additional control (moderator) analyses with reduced publication bias were conducted, in which the residuals from PEESE (i.e., derived observed Hedge's g minus expected Hedge's g based on PEESE) were used as dependent variable. Outliers identified by the sensitivity analysis were eliminated from the analysis[3].

3. Results

3.1. Systematic review of how the TSST was implemented.

A total of N = 143 studies were included into this systematic review. Most of them (n=130, 90.9%) used either the TSST-C, the TSST-M or the original TSST (see Table 2)[4]. Due to this overall preference in the TSST versions used, a detailed description of the results by version is provided in the supplementary materials (Tables S1 – S5).

3.1.1 Study details: Age range, time, exclusion criteria

Age ranges. A total of 15.4% of the studies (n=22; 72.7% using the TSST-C) did reported no age range required for participation. Studies that did provide specific age ranges used the TSST-C (total range 6-18 years, see figure S1a) mostly in children samples only (34%) and mixed samples with children and adolescents (35.1%), while the TSST-M (total range 7-17 years, see Figure S1b)

[3] This differs from the approach outlined in the preregistration and is rooted in extreme z-values of the identified outliers.

[4] Studies were categorized by the TSST version and additional information provided in the Methods sections. If the information was insufficient for categorization, the study was categorized as "residual".

was mainly used in mixed samples (47.8%), and the TSST (total range 8-18 years old, see figure S1c) for adolescents between 12 and 18 years of age (69.2%).

Table 2. Overview of the included studies and their respective effect sizes according to the TSST version used.

Version of the TSST	Publications used in qualitative analysis		Publications and effect sizes used in quantitative analysis[5]			
	N=143		*N*=91		*K*=121	
	n	*%*	*n*	*%*	*k*	*%*
TSST-C (Buske-Kirschbaum et al., 1997)	94	65.7	59	64.8	79	65.3
TSST-M (Yim et al., 2010)	23	16.1	13	14.3	18	14.9
TSST (Kirschbaum et al., 1993)	13	9.1	11	12.1	13	10.7
Modified TSST (Kirschbaum et al., 1993)	3	2.1	1	1.1	2	1.7
TSST-T (Cameron et al., 2017)	3	2.1	3	3.3	4	3.3
TSST-G (von Dawans et al., 2011)	1	0.7	0	0	0	0
TSST-VR (Kelly et al., 2007)	1	0.7	1	1.1	1	0.8
TSST-OL (Gunnar et al., 2021)	1	0.7	1	1.1	1	0.8
Other TSST modifications	4	2.8	2	2.2	3	2.5

Time windows. 23.8% of all studies (*n*=34; 58.8% using the TSST-C) failed to state a time window for TSST administration sessions. Most studies using the TSST-C (29.8%) scheduled appointments between 2pm and 5pm (details see Figure S2a), studies using the TSST-M favored 12pm to 5pm (34.8%; details see Figure S2b), and those using the TSST used slots between 2pm and 5pm (38.5%; details see Figure S2c). Only three studies reported time windows before 12pm.

Exclusion criteria. 74.13% of studies (*n*=106) specified exclusion criteria; 105 (99.1%) thereof relied on participant-centered criteria and seven (6.6%) on parent-centered criteria only or in addition to the participant-centered ones. For detailed quotas according to the TSST used, see Table S1.

Participant-centered criteria. The most often applied exclusion factor was the use of an HPA-affecting medication (58.1%), with most studies specifically naming steroid medication (37.1%). The second most frequent exclusion factor was mental health issues (54.3%) with most naming acute and chronic health problems, such as an acute or life-time diagnosis of depression.

[5] After correction for publication bias.

As the third were physical health issues (53.3%), with most studies focusing on chronic illnesses such as diabetes (46.7%). Developmental impairments were a criterion in 30.5% of cases, with most (18.1%) focusing on general developmental disorders such as learning disabilities, while cognitive impairments were used in 29.5% of cases with most (16.2%) not specifying an IQ threshold; with a specified threshold, most required an IQ of at least 80 (7.6%). 16.2% of studies excluded participants suffering from neurological disorders such as epilepsy, and 14.3% did so on grounds of insufficient language skills in the language used in the experiment. Only 4.8% of studies excluded participants with sensory impairments (e.g., vision, hearing, or speech impairments) or participants living with psychosocial problems such as (past) abuse in family life (3.8%). Other reasons for exclusion included participant sex (1%), current pregnancy (5.7%), stages of puberty (5.7%) and complications during a participant's birth (3.8%).

Parent-centered criteria. Studies excluded participants whose parents reported drug abuse during pregnancy (42.9%), mental health problems (28.6%) or chronic physical health problems (14.3%). Two studies using the TSST-C excluded participants due to parent's insufficient language skills, mainly due to their inability to fill out the questionnaires.

3.1.2 TSST details: Duration, timing, judges, experimenter, and saliva collection

Phase durations report. Overall, 111 studies (77.6%) systematically delivered information on the duration of both the acclimation time before starting the TSST and the TSST phases individually. Fourteen studies (9.8%) did not report on the acclimation time while reporting information on the TSST phases. 13 papers (9.1%) did provide the acclimation time, but not on (all) the TSST phases with most (n=10) not including any phase durations, two not mentioning the preparation phase and one not reporting the duration of any task. Finally, five papers (3.5%) failed to report on either acclimation time or phase durations. A Pearson's Chi-Squared test showed no significant evidence for an association between the TSST version used and information reported on acclimation time and duration of TSST phases ($\chi 2(21)=18.61$, $p=.610$).

Phase durations. On average, participants became acclimated to the experimental setting for 32.66 ± 19.92 minutes (n=124; 0 – 120 minutes) and had 5.15 ± 1.93 minutes (n=126; 2 – 10 minutes) of preparation time before performing for 4.91 ± 0.59 minutes (n=128; 2 – 6 minutes) in the public speaking task and 4.80 ± 0.61 minutes (n=128; 1 – 5 minutes) in the arithmetic task with the entire TSST taking 14.86 ± 2.28 minutes (n=125; 6 – 20 minutes). Comparing the means between TSST-C, TSST-M and the TSST (see Table S2), the Kruskal-Wallis rank sum test showed no significant differences in the duration of the preparation phase, H(2)=4.73, p=.094, or in the TSST's total duration , H(2)=2.86, p=.238. However, we detected significant differences in the acclimation time, H(2)=6.86, p=.032, f=.22, duration of the public speaking task, H(2)=21.50, p<.0001, f=.45, and duration of the arithmetic task, H(2)=10.62, p<.01, f=.28. Post-hoc group comparisons were performed using Dunn-Bonferroni's tests. Regarding the acclimation time, we noted a significant difference with the TSST-C needing a longer acclimation time on average than

the TSST with a small to medium effect (z=-2.59, adj. p=.029, r=.27). Considering the public speaking task's duration, we detected significant differences with the TSST-M requiring a longer task on average than the TSST-C with a medium to large effect (z=4.11, adj. p<.0001, r=.40) and the TSST with a large effect (z=-3.99, adj. p<.0001, r=.69). Lastly, concerning the arithmetic task's duration, we identified a significant difference with the TSST-M having an on average shorter duration than the TSST-C with a medium to large effect (z=-3.18, adj. p<.01, r=-.31). For detailed quotas by used TSST version, see Table S2.

Task explanation timing. Most studies (95.8%) did not specify when the TSST task was first explained to the participant (e.g., start of the appointment vs. before the preparation phase). Only six studies (n=5 using the TSST-C) specified that that occurred right before the preparation phase.

Details about jury. One hundred and sixteen (81.1%) studies reported on the number of judges (79.3% using two judges). 21 studies reported on the judges-to-participant gender ratio in general with most (76.2%) reporting mixed-gender juries. 59 studies reported on non-verbal feedback given by the jury, and 91.5% of these described neutral behavior. For details according to the TSST version, see Table S2.

Details about experimenter. Most studies did not provide details about the experimenter leading the session. Only three studies using the TSST-M stated that they always had female presenting experimenters, while seven studies reported on the experimenter's neutral interaction behavior towards the participant. Overall, 53 studies clarified whether the experimenter was on the jury with most (81.1%) stating clearly that they were not. For details by TSST version, see Table S2.

Saliva collection. 126 publications reported on the saliva sampling method with most studies (51.8%) using Salivettes (Sarstedt, Germany), 51 studies (35.7%) using passive drool collection, and only one study using SaliCaps. For details by TSST version, see Table S3.

3.1.3 Descriptive data

Responder rates. Ten studies in total reported cortisol responder rates, of which 33.3% (three using the TSST-C) used the 10%-criterion defined by Ji et al. (2016); see Table S4 and Table S5). Overall, responder rates between 59% and 82% were reported for healthy control samples with responder rates falling with the sample's size

Mean baseline and peak cortisol data. The median reported baseline salivary cortisol concentration was 3.06nmol/L (*IQR*=1.07, 5.05) measured at a median baseline time of -0.5 minute (*IQR*=-2.5, 1.5) relative to the start of the TSST preparation phase. The median peak salivary cortisol concentration was 5.61nmol/L (*IQR*=1.45, 9.77) measured at a median peak time of 25 minutes (*IQR*=10, 35) relative to the start of the TSST preparation phase.

3.2. Baseline to peak differences in salivary cortisol response – a meta-analysis

In total, 64.3% of the studies used in our qualitative analysis (n=92) with a total sample of N=8291 participants (N_{age}=6902; M_{age}=11.95, SD=2.22) were included in the quantitative analysis. Three thousand five hundred and fifty-six participants were children (42.9%), and 4588 participants were adolescents (55.3%); 147 participants could not be assigned to a specific age group (1.8%). Most belonged to mixed-gender samples (n=6199; 74.8%), while 976 (11.8%) participants came from all-female groups, 874 (10.5%) came from all-male groups, and 242 (2.9%) participants could not be assigned to a gender grouping. Relying on these data, we calculated 123 effect sizes for the change in salivary cortisol concentrations.

Comparing the included publications (see Table S6) between this meta-analysis and Seddon et al. (2020), 39 publications (42.4%) were new due to being published after February 2018. Thirty-three publications included in Seddon et al. (2020) were not included in the present analysis because they: 1) did not report cortisol data (k=9, 27.3%), 2) used a TSST that was significantly altered compared to the standard protocol, 3) were classed as gray literature, 4) failed to supply raw data (each k=7, 21.2%), 4) reported only on clinical samples (k=2, 6.1%) or 5) did not include the TSST's arithmetic task (k=1, 3%).

3.2.1 Mean effect size

Testing overall homogeneity, our results suggest that it is unlikely that effect sizes originated from the same population (Q(122)=927.58, p<.0001, I^2=90.8%). Indeed, the between-study variance (σ^2=.24) was substantial, thus justifying that we employed random-effects modeling to analyze the data. Results of the mixed effects model showed the TSST's medium effect on the average baseline-to-peak salivary cortisol response (k=123; g=.61, z=12.84, p<.0001). The skewness of the effect sizes measured 2.93 with a positive excess kurtosis of 14.85, indicating strongly left-skewed normal distribution (see Fig. 2).

3.2.2 Publication bias assessment

Publication bias was assessed applying three methods: meta-regression using the year of publication as a moderator, the Egger's regression asymmetry test, and PET-PEESE[6].

Meta-regression showed no significant evidence for bias based on publication year (k=123, z_{year}=-1.69, p=.091; see Figure S4a). Sensitivity analysis indicated two extreme outliers (k=123, z=12.84, p<.0001; see Figure S5a). Removing these two outliers from the meta-regression resulted in a significant influence of the publication year (k=121, z_{year}=-3.11, p<.01) with earlier studies reporting higher effect sizes (see Figures S4b and S5b). Egger's test (k=121, z=5.36, p<.0001) and

[6] Due to necessity, the presented analysis plan uploaded to OSF was adapted.

PET-PEESE consistently yielded significant evidence for publication bias (PET: k=121; β_{se}=2.37, p<.0001; PEESE: β_{vg}=5.47, p<.0001). Hence, in addition to Hedge's g, the PEESE residuals were used as dependent variables in our control moderator analyses (see Table 4).

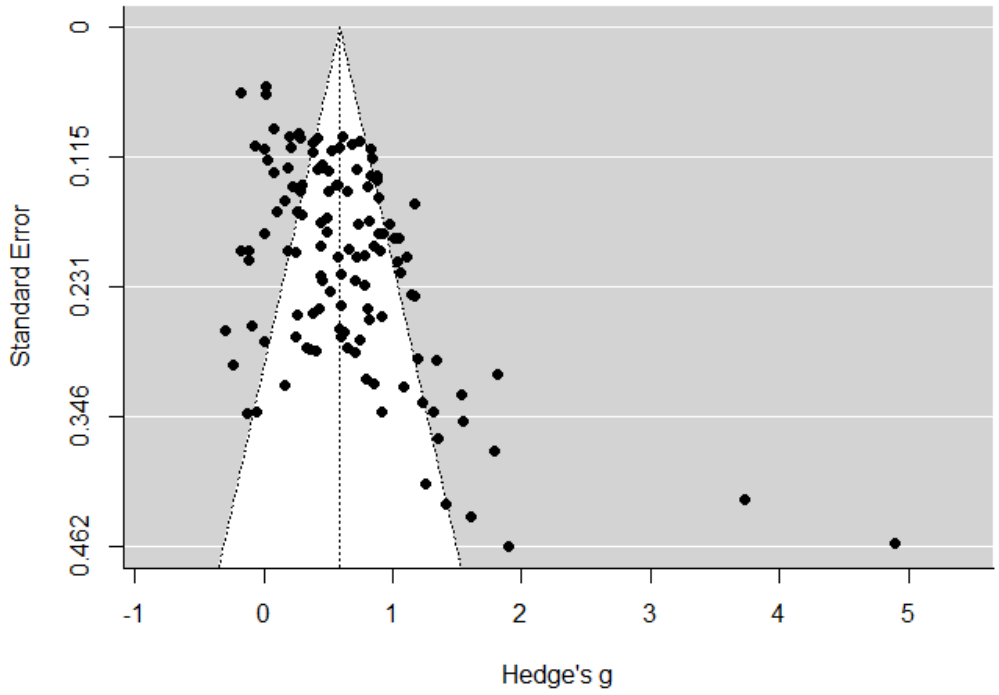


Figure 2. Funnel plot of effect sizes before the correction of the publication bias.

Regarding the true effect size estimate, PET intercept was non-significant (k=121; $\beta_{0_{PET}}$=.10, p=.241), indicating no effect of the TSST after removing publication bias. In contrast, PEESE intercept still indicated a significant "true" effect of the TSST on salivary cortisol after removing publication bias (k=121; $\beta_{0_{PEESE}}$=.32, p<.0001). While conventions suggest the use of PET estimators if the PET intercept is not significantly different from 0 (Carter et al., 2019), we note that 90.2% of the reported effect sizes in this analysis were larger than zero (see Figure S3). Thus, while PET may yield overcorrected estimators (as also noted by Aksayli et al., 2019), PEESE may be considered to offer more reliable estimators for the present data. This suggests a moderate (publication bias corrected) effect regarding the baseline-to-peak cortisol response to the TSST (k=121; g =.56, z=14.96, p<.0001).

3.3 Moderator analyses

Due to a lack in variance in the reported data on the genders of judges and experimenter, as well as experimenter behavior, we were unable to run those planned analyses.

3.3.1 Confirmatory analyses

Of all intended confirmatory analyses, the participant sample gender, number of judges and the TSST's total duration yielded statistically significant results, which are discussed here. A list of results for categorical variables is summarized in Table 4, for results for continuous variables see Table 5.

Sample gender. Sample genders varied between studies from all male to all female, and effect sizes without gender ratio information were cut from the analysis. The moderating effect of the sample gender class was not significant for Hedge's *g*, Q(2)=5.58, *p*=.061, but for PEESE residuals, Q(2)=6.29, *p*=0.043, as dependent variables. Since the PEESE calculation of residuals aims to remove publication bias, their results were estimated to be more reliable, hence indicating stable moderation. Model results with planned contrasts showed that all gender classes differed significantly from 0 with the "all-female" samples having an on average significantly bigger effect than the "all-male" samples, Q(1)=5.220, estimate=.287, se=.126, z=2.285, p=.022.

Number of judges. Jury size varied among studies from one to three individuals, but only the factor levels "1 judge" and "2 judges" were entered into analysis. The moderating effect of this number was significant for both Hedge's *g*, Q(1)=11.23, *p*<.0001, and PEESE residuals, Q(1)=13.23, *p*=0.0003, as dependent variables, indicating stable results beyond publication bias. Model results revealed that only the "2 judges" factor level's effect was significantly different from 0, *β*=.5675. *se*=.0511, *z*=11.1155, *p* < .0001, while the "1 judge" factor level was not significantly different, *β*=.0314. *se*=.1516, *z*=.2068, *p*=.836.

Total duration of the TSST. The total TSST time varied from 6 to 20 minutes and was entered as a continuous variable. The moderating effect was significant for both Hedge's *g*, Q(1)=11.96, *p*=0.0005, and PEESE residuals, Q(1)=8.33, *p*=0.0039, as dependent variables, indicating stable results beyond publication bias. Model results revealed a small positive effect, *β*=.0626. *se*=.0181, *z*=3.459, *p*=.0005, indicating larger effect sizes with longer durations.

3.3.2 Exploratory analyses

Moderation of effect size by baseline cortisol. The cortisol awakening response (CAR; Pruessner et al., 1997) is a well-established circadian pattern defined by a spike in cortisol level after the first awakening in the morning and a slow fall throughout the day. This variation in readily available cortisol in participants (see Adam et al., 2006) might in turn systematically affect and moderate the cortisol reactivity to the TSST (see Kudielka et al., 2004; Schommer et al., 2003),

which can be tested by either focusing on the administration time or the measured baseline cortisol itself. In their meta-analysis, Goodman et al. (2017) tested this potential first moderation and recommended that TSSTs be administered in the afternoon between 2pm and 5pm to ensure maximum pre-to-post TSST cortisol reactivity. In an attempt to replicate their results, we divided the studies in this review into "early" (before 12pm, k=4), "midday" (between 12pm and <2pm, k=30), and "late" (after 2pm, k=62) starters by the administration window start and entered into a moderation analysis. This model, however, did not reach significance, Q(2)=2.96, p=.228. A second model that combined the first two classes into an analysis "early" (k=34) vs. "late" (k=62) starters to even out the group sizes, did not reach significance either, Q(1)=.69, p=.405. Since this could be due to still very dissimilar group sizes, we decided to test the second moderation model with the measured baseline cortisol itself.

An initial moderation model including all effect sizes (k=121) did not reach significance, Q(1)=3.36, p=0.067. Descriptively though, we noted a trend toward a negative effect of higher baseline cortisol, β=-.038. se=.0209, z=-1.8334. A bubble plot further revealed potential extreme outliers in baseline cortisol (see Figure S6a) that might have suppressed a significant moderation effect. To remove extreme outliers, a sample size weighed the average of the baseline cortisol was calculated (M=4.057, SD=1.836) and used to define a cut-off of two standard deviations over and under the mean. Thus, all effect sizes produced by baseline cortisol measurements in the average range (0.383nmol/l - 7.729nmol/l, k=116) were entered into the second moderation analysis. This new model was significant, Q(1)=5.998, p=0.014. Model results revealed a small negative effect, β=-.0681. se=.0278, z=-2.4491, [-.1226, -.0136], suggesting higher baseline cortisol producing lower effect sizes pre-to-post TSST (see Figure S6b).

TSST version and age class interaction. The TSST-C and TSST-M were validated with participants of similar sample age groups and differ predominantly in the public speaking task's topic (see Table 1). We therefore posed the question whether there was any difference in effect sizes between these two TSST versions in interaction with the population age groups. This interaction was tested in two separate models, one with effect sizes derived from studies using the TSST-C only, and the other using the TSST-M only. The TSST-C and sample age group interaction was not significant, Q(1)=2.10, p=.147, indicating no difference. With the TSST-M however, the interaction was significant, Q(1)=5.07, p=.024, indicating significant differences in effect sizes between the sample age groups. These model results revealed that the TSST-M's effect was significantly larger in the child samples, β=.6361. se=.1211, z=5.2545, p<.0001, than in the adolescent samples, β=.3155. se=.1071, z=2.9456, p=.0032. For detailed results, see Table 4.

Table 4. Results of moderation analyses using categorical moderators.

		Hedge's *g*					PEESE	
Moderators in confirmatory analysis	*k*	*G*	95% CI	Q_{model}	*df*	*p*	*Q*	*p*
TSST version	110			1.427	2	.490	1.151	.562
TSST-C	79	.599	[.493, .705]			<.0001		
TSST-M	18	.459	[.242, .677]			<.0001		
TSST	13	.519	[.269, .768]			<.0001		
residual group	11							
Sample age	116			.088	1	.767	.290	.590
Children	52	.560	[.441, .679]			<.0001		
Adolescents	64	.538	[.433, .642]			<.0001		
not specified	5							
Sample gender	116			5.584	2	.061	6.290	.043
Only males	13	.439	[.265, .612]			<.0001		
Mixed gender	86	.541	[.448, .634]			<.0001		
Only females	17	.590	[.410, .770]			<.0001		
not specified	5							
Judges gender	16			.003	1	.958	.607	.436
All female	4	.735	[.377, 1.092]			<.0001		
Mixed	12	.723	[.472, .974]			<.0001		
Opposite to participant	3							
not specified	102							
Judges number				11.229	1	<.001	13.229	<.001
One	9	.031	[-.266, .329]			.836		
Two	71	.568	[.468, .668]			<.0001		
Three	2							
not specified	38							
Judges feedback	52			1.433	1	.231	.020	.887
Neutral	48	.481	[.344, .618]			<.0001		
Friendly	4	.801	[.295, 1.307]			.002		
Progressively negative	1							
not specified	68							
Acclimation time				4.843	4	.304	5.250	.263
0 – 15 minutes	17	.617	[.392, .842]			<.0001		
16 – 30 minutes	52	.475	[.344, .606]			<.0001		
31 – 45 minutes	16	.526	[.312, .740]			<.0001		
46 – 60 minutes	16	.742	[.522, .962]			<.0001		
More than 60 minutes	6	.664	[.254, 1.074]			.002		
not specified	14							
Moderators in exploratory analysis	*k*	*G*	95% CI	Q_{model}	*df*	*P*		
Age * TSST-C	74			2.099	1	.147		
Children	43	.545	[.417, .674]			<.0001		
Adolescents	31	.678	[.534, .823]			<.0001		
not specified	5							
Age * TSST-M	18			5.070	1	.024		
Children	9	.636	[.399, .873]			<.0001		
Adolescents	9	.316	[.106, .526]			.003		
Administration time start	96			2.960	2	.228		
Early (<12pm)	4	.223	[-.286, .732]			.390		
Midday (>12pm, <2pm)	30	.632	[.472, .793]			<.0001		
Late (>2pm)	62	.516	[.408, .623]			<.0001		
not specified	25							

Note. * < .05. ** < .01. *** < .001. Moderators with less than 4 effect sizes were not tested, factor levels with *k*=0 are not listed.

Table 5. Results of moderation analyses using continuous moderators.

		Hedge's *g*				PEESE	
Moderators in confirmatory analysis	*K*	*g*	95% CI	Q_{model}	*p*	*Q*	*p*
Linear model of sample age	96			.541	.462	1.825	.177
Intercept		.712	[.242, 1.183]		.003		
Sample age		-.014	[-.052, .024]		.462		
Linear model of acclimation time	107			1.231	.267	1.732	.188
Intercept		.476	[.305, .647]		<.0001		
Acclimation time		.002	[-.002, .007]		.267		
Linear model of total TSST duration	81			11.963	<.001	8.329	.004
Intercept		-.380	[-.902, .143]		.154		
Total TSST duration time		.063	[.027, .098]		<.001		
Moderators in exploratory analysis	*K*	*g*	95% CI	Q_{model}	*p*		
Linear model of baseline cortisol (model 1)	121			3.361	.067		
Intercept		.689	[.513, .864]		<.0001		
Cortisol at baseline		-.038	[-.079, .003]		.067		
Linear model of baseline cortisol (model 2)	116			5.998	.014		
Intercept		.781	[.573, .989]		<.0001		
Cortisol at baseline		-.068	[-.123, -.014]		.014		

Note. * < .05. ** < .01. *** < .001. Moderators with less than 4 effect sizes were not tested, factor levels with *k*=0 are not listed.

4. Discussion

The present work investigated the experimental stress induction in children and adolescents with the Trier Social Stress Test with a special focus on the modulation by specific protocol variations.

Our systematic review of 143 studies identified three predominantly used TSST versions: the TSST-C (65.7%) and TSST-M (16.1%) specifically adapted for the full youth age spectrum, and the original TSST (9.1%) employed in adolescents to young adult samples. TSST tasks included in the TSST-M were significantly shorter than in the other two versions. Most of the studies employed a mixed-gender panel with two members usually giving non-verbal, neutral feedback. A majority of studies reported excluding participants presenting factors known to affect HPA functioning such as endocrinologically relevant medications, acute/chronic mental and physical health issues (each in more than 50% of studies). Notably, a significant amount of methodological details (up to 90% of data per observed variable) could not be extracted due to limited information in the publications.

Our meta-analysis summarizing 91 studies (*k*=121, *N*=8177) with mainly mixed-gendered and mixed-aged samples revealed a medium overall baseline-to-peak effect (*Hedge's g* = 0.56), which is in line with the overall effect reported by Seddon et al. (2020; *Cohen's d* = 0.47) indicating that the TSST's youth-specific adaptations are effective tools to elicit salivary cortisol responses in children and adolescents in general. We identified no significant differences between the TSST-C and TSST-M in terms of cortisol reactivity. However, the number of judges, gender of participants, the total duration of the TSST and baseline cortisol levels, as well as the interaction

of age groups and TSST versions turned out to be significant moderators, all of which are discussed below.

Our first major finding is the observation that youth samples exhibit a less pronounced cortisol stress reaction to the TSST than do adult samples, which is consistent with previous reports. Seddon et al. (2020) also reported on medium pre-to-post TSST effect sizes in youth samples, while adult samples showed large effect sizes, estimated consistently at $d = 0.925$ (Goodman et al., 2017) and $d = 0.930$ (Gu et al., 2022) in meta-analyses. What are the possible explanations for this reduced cortisol responsiveness in children and adolescents compared to adults?

First, children and adolescents might perceive the TSST's specific adaptations as less stressful and thus tend to exhibit a blunted stress response to this laboratory stressor. This assumption is supported by the fact that in some studies with younger children, the aspect of social threat in particular (e.g., by refusing verbal/non-verbal feedback) is made milder to keep the stress in the participating children at a tolerable and ethically acceptable level. An indirect indication of the presumed effect of lower social threat in the children's TSST can perhaps be derived from the observation that the cortisol response is significantly stronger when at least two people are present in the panel. Unfortunately, the statistical power here is limited, as there are very few studies involving a single-person panel. It is also possible that the experimental procedures were terminated more frequently in children due to particularly severe stress reactions in children, although this remains speculative and cannot be conclusively assessed due to the lack of usable data from both children/adolescents and adults.

The second explanation is based on the assumption that the HPA axis matures during childhood and adolescence. There is some supporting evidence for this, suggesting that the HPA axis attains its maximum responsiveness in adolescence or young adulthood. However, within the age range from 9 to 18 years, the results appear paradoxical. Seddon et al. (2022) reported stronger cortisol responses in children than adolescents. In contrast to this, Gunnar et al. (2009) directly compared age groups in their single study and reported that adolescents (13–15-year-olds) tended to show a stronger cortisol response to the TSST than children (9–11-year-olds). Another study reported the same pattern, whereby the effect was particularly pronounced in performance-related tasks (Stroud et al., 2009). In the present meta-analysis, which relies on a significantly larger number of effect sizes, we were able to confirm this effect in neither one direction nor the other. A significant age effect was detected only in conjunction with the TSST-M, a finding in line with Seddon et al. (2020) that might be attributable to the wider age range of participants in those studies. Since acute cortisol responsiveness can hardly be compared because of the significant differences between the stressors used in child/adolescent studies and adult studies, this question will remain unanswered as long as longitudinal studies are lacking on the development of HPA axis reactivity using experimental paradigms.

Aside from this, developmental interactions with other stress-responsive hormonal systems such as gonadal steroids (Domes et al., 2024), accumulating experiences with TSST-like everyday

situations (Wüst et al., 2005; Batabyal et al., 2021; Barthel et al., 2024), age-related differences in stress coping styles as a trait (Illius et al., 2024) and in stress management strategies as a state (Goessl et al., 2017; Thabrew et al., 2018), daily life stressors (Weber et al., 2022), and the socioeconomic and health status of the family (Lupien et al., 2000) are some of the known influence factors that might account at least partially for the reduced cortisol responsiveness in children and adolescents, either by themselves or in interaction with each other.

Some of the moderators identified in previous meta-analyses were confirmed in our analysis. (1) *Sample gender* was a significant moderator in our analysis, but the direction of the effect was heterogenous: while Seddon et al. (2020) reported higher cortisol responses in mixed-gender samples compared to exclusively male and female samples; our analysis suggests that girls exhibit the highest responsiveness to the TSST. This effect contrasts with studies in adults indicating that men display larger effects than women (Gu et al. 2022). At this point, it seems plausible that the fluctuation of sex steroids in pubertal girls (Ordaz and Luna, 2012) and the increased cortisol binding capacity in women using oral contraceptives (Liu et al., 2017) could explain these effects.

(2) As in previous studies in adults (Goodman et al. 2017), the *number of judges* was a significant moderator: More than one judge resulted in significantly larger cortisol reactions in both youth and adult samples, which might be simply due to the higher level of social threat and uncontrollability induced by two or more panel members.

(3) In addition, *neutral feedback* in the sense of social unresponsiveness seemed to trigger a stronger cortisol response than feedback with emotional valence. However, we observed no moderation by *judges' gender*, in contrast to previous findings in both children/adolescents and adults, which reported stronger cortisol responses to mixed-gender than to all-female panels (Goodman et al., 2017, Seddon et al., 2020).

(4) For protocol variations, the *TSST*'s *duration* was a significant moderator in both this meta-analysis and Seddon et al. (2020)'s analysis with longer durations resulting in higher effect sizes, which might reflect a stress "dosage" effect. Although other stressors are much shorter combining a pain stimulus with social evaluation (eg, the socially evaluated cold pressor test, SECPT; Schwabe et al. 2008), a minimal TSST duration might be needed to activate the HPA axis. Although our analysis did not confirm that the *time of day* at which the TSST was carried out (morning vs. afternoon) was a significant moderator, the trend we observed is in line with Seddon et al. (2020) who reported larger effects in the afternoon. This may be largely due to the fact that cortisol levels gradually decline during the day, with low levels in the later afternoon, potentially leaving more room for pronounced reactions to acute challenges. In line with this assumption, our exploratory analysis suggested that low baseline cortisol levels predicted higher cortisol responses to the TSST.

Some limitations of the present study need to be considered. Firstly, we were unable to find a significant amount of data in the original publications (15-90% of observable instances, depending on the observed variable) extending across all areas of interest in our systematic review and – data that could not be extracted for this qualitative description and quantitative meta-analysis. We were therefore unable to carry out all the analyses we had planned and those we did perform sometimes have low statistical power. In line with this, over a third of the papers identified in our systematic review were excluded from our meta-analysis because of missing cortisol data. This, too, represents a significant loss of information that could not affect both moderator analyses as well as a general estimation of global effects and publication bias. Thirdly, note that we omitted unpublished data. Finally, we focused on studies with unselected samples or healthy samples to provide estimates as unaffected as possible by clinical or subclinical condition. We are aware that this decision reduces the variability within the dataset and presumably compromises the generalizability of our analysis. A follow-up analysis could explicitly include different clinical samples and compare the effects seen between disorders and levels of symptom severity.

Two implications should also be mentioned: The first one is the lack of methodological details in many original studies, which hindered a full-scale moderator analysis in this and previous meta-analyses of all potentially available data. In our opinion, this status quo reveals the need for consensus guidelines, and reporting standards for future studies (cf. Stalder et al., 2022, for the cortisol awakening response) and open data to enable qualitative and quantitative syntheses in the future. This is in line with the general desire to provide full methodological details and primary data for secondary research synthesis and to at least partially overcome the replication crises (Narvaez Linares et al., 2020; Pennington, 2023).

The second implication refers to the weaker overall reactivity to be expected from the TSST in youth samples than in adult samples. This should be considered when classifying individuals into “responders” and “non-responders” based on their peak-to-baseline responses. The traditionally used hard baseline-to-peak cut-offs of 1.5nmol/l or 2.5nmol/l might not be suitable in children and adolescents. The 10% increase-criterion defined by Ji et al. (2016) and the 15.5% increase-criterion by Miller et al. (2013) might be more suitable, however their validity in children and adolescents still need to be tested.

In conclusion: the present meta-analysis together with previous studies show that the TSST with its specific adaptations for children and adolescent is an effective tool for inducing cortisol stress responses in a youth sample. The cumulative evidence in children/adolescents (TSST-C and TSST-M) and adults (TSST) suggests that the protocol is most effective when it is performed as intended by the original design: in full length with at least two mixed-gendered judges giving neutral, non-verbal feedback to the participant – even in studies involving the TSST’s specific adaptations for children and adolescents. On the other hand, the data available to date provide no indication that it makes a difference which of the two main variants for children (TSST-C vs. TSST-M) is used.

Declaration of Competing Interest

None.

Funding

This work was supported by the University Research Priority Program “Psychobiology of Stress” funded by the State Rhineland-Palatinate.

CRediT authorship contribution statement

Saskia Seel: Conceptualization, Screening, Data collection, Investigation, Data curation, Project administration, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Bernhard Pastötter:** Formal analysis, Writing – review & editing. **Gregor Domes:** Conceptualization, Writing - review & editing, Data curation, Project administration, Methodology.

Acknowledgements

The review protocol, data, codebook, and analysis code related to this paper are available at OSF: https://osf.io/4w9s7/. This work was not preregistered.

References

A list of all references used in the systematic review and meta-analysis can be found in the Supplementary to this work.

Adam, E. K., Hawkley, L. C., Kudielka, B. M., Cacioppo, J. T., 2006. Day-to-day dynamics of experience–cortisol associations in a population-based sample of older adults. Proceedings of the National Academy of Sciences, 103(45), 17058-17063.

Aksayli, N. D., Sala, G., Gobet, F., 2019. The cognitive and academic benefits of Cogmed: A meta-analysis. Educational Research Review, 27, 229-243.

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. https://doi.org/10.1016/j.neubiorev.2013.11.005

Allen, A.P., Kennedy, P.J., Dockray, S., Cryan, J.F., Dinan, T.G., Clarke, G., 2017. The trier social stress test: principles and practice. Neurobiol. Stress 6, 113–126. https://doi.org/10.1016/j.ynstr.2016.11.001

Barthel, M.-C., Fricke, K.L., Muehlhan, M., Vogel, S., Alexander, N., 2024. Habituation of the biological response to repeated psychosocial stress: a systematic review and meta-analysis [Poster presentation]. Presented at the Psychologie und Gehirn 2024, Hamburg.

Batabyal, A., Bhattacharya, A., Thaker, M., Mukherjee, S., 2021. A longitudinal study of perceived stress and cortisol responses in an undergraduate student population from India. PLOS ONE 16, e0252579. https://doi.org/10.1371/journal.pone.0252579

Bengtsson, H. (2023). matrixStats: Functions that Apply to Rows and Columns of Matrices (and to Vectors). R package version 1.2.0, URL: https://CRAN.R-project.org/package=matrixStats .

Borenstein, M., Hedges, V. V., Higgins, J. P. T., Rothstein, H. R., 2011. Introduction to Meta-Analysis. John Wiley & Sons, Ltd. DOI:10.1002/9780470743386

Brand, S., Wilhelm, F.H., Kossowsky, J., Holsboer-Trachsler, E., Schneider, S., 2011. Children suffering from separation anxiety disorder (SAD) show increased HPA axis activity compared to healthy controls. J. Psychiatr. Res. 45, 452–459. https://doi.org/10.1016/j.jpsychires.2010.08.014

[1]Cameron, C.A., McKay, S., Susman, E.J., Wynne-Edwards, K., Wright, J.M., Weinberg, J., 2017. Cortisol Stress Response Variability in Early Adolescence: Attachment, Affect and Sex. J. Youth Adolesc. 46, 104–120. https://doi.org/10.1007/s10964-016-0548-5

Canty, A., & Ripley, B. (2022). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28.1.

Carter, E.C., Schönbrodt, F.D., Gervais, W.M., Hilgard, J., 2019. Correcting for Bias in Psychology: A Comparison of Meta-Analytic Methods. Adv. Methods Pract. Psychol. Sci. 2, 115–144. https://doi.org/10.1177/2515245919847196

Champely, S. (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0, URL: https://CRAN.R-project.org/package=pwr .

Childs, E., Vicini, L.M., De Wit, H., 2006. Responses to the Trier Social Stress Test (TSST) in single versus grouped participants. Psychophysiology 43, 366–371. https://doi.org/10.1111/j.1469-8986.2006.00414.x

Coburn, K. M., & Vevea, J. L. (2019). weightr: Estimating Weight-Function Models for Publication Bias. R package version 2.0.2, URL: https://CRAN.R-project.org/package=weightr .

de Veld, Danielle M. J., Riksen-Walraven, J.M., de Weerth, C., 2014a. Acute psychosocial stress and children's memory. Stress Amst. Neth. 17, 305–313. https://doi.org/10.3109/10253890.2014.919446

de Veld, Danielle M. J., Riksen-Walraven, J.M., de Weerth, C., 2014b. Does the arrival index predict physiological stress reactivity in children. Stress Amst. Neth. 17, 383–388. https://doi.org/10.3109/10253890.2014.936004

Dickerson, S.S., Kemeny, M.E., 2004. Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. Psychol. Bull. 130, 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Domes, G., Linnig, K., Von Dawans, B., 2024. Gonads under stress: A systematic review and meta-analysis on the effects of acute psychosocial stress on gonadal steroids secretion in humans. Psychoneuroendocrinology 164, 107004. https://doi.org/10.1016/j.psyneuen.2024.107004

Fisher, Z., Tipton, E., & Zhipeng, H. (2023). robumeta: Robust Variance Meta-Regression. R package version 2.1, URL: https://CRAN.R-project.org/package=robumeta .

Gartlehner, G., Hansen, R.A., Morgan, L.C., Thaler, K., Lux, L., Van Noord, M., Mager, U., Thieda, P., Gaynes, B.N., Wilkins, T., Strobelberger, M., Lloyd, S., Reichenpfader, U., Lohr, K.N., 2011. Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. Ann. Intern. Med. 155, 772–785. https://doi.org/10.7326/0003-4819-155-11-201112060-00009

Goessl, V.C., Curtiss, J.E., Hofmann, S.G., 2017. The effect of heart rate variability biofeedback training on stress and anxiety: a meta-analysis. Psychol. Med. 47, 2578–2586. https://doi.org/10.1017/S0033291717001003

Goodman, W.K., Janson, J., Wolf, J.M., 2017. Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test. Psychoneuroendocrinology 80, 26–35. https://doi.org/10.1016/j.psyneuen.2017.02.030

Gordis, E.B., Granger, D.A., Susman, E.J., Trickett, P.K., 2006. Asymmetry between salivary cortisol and alpha-amylase reactivity to stress: relation to aggressive behavior in adolescents. Psychoneuroendocrinology 31, 976–987. https://doi.org/10.1016/j.psyneuen.2006.05.010

Gu, H., Ma, X., Zhao, J., Liu, C., 2022. A meta-analysis of salivary cortisol responses in the Trier Social Stress Test to evaluate the effects of speech topics, sex, and sample size. Compr. Psychoneuroendocrinology 10, 100125. https://doi.org/10.1016/j.cpnec.2022.100125

Gunnar, M. R., Frenn, K., Wewerka, S. S., Van Ryzin, M. J., 2009. Moderate versus severe early life stress: Associations with stress reactivity and regulation in 10-12-year-old children. Psychoneuroendocrinology, 34(1), 62–75. https://doi.org/10.1016/j.psyneuen.2008.08.013

Gunnar, M.R., Reid, B.M., Donzella, B., Miller, Z.R., Gardow, S., Tsakonas, N.C., Thomas, K.M., DeJoseph, M., Bendezú, J.J., 2021. Validation of an online version of the Trier Social Stress Test in a study of adolescents. Psychoneuroendocrinology 125, 105111. https://doi.org/10.1016/j.psyneuen.2020.105111

Hellhammer, D.H., Wüst, S., Kudielka, B.M., 2009. Salivary cortisol as a biomarker in stress research. Psychoneuroendocrinology 34, 163–171. https://doi.org/10.1016/j.psyneuen.2008.10.026

Higgins, J. P. T., Green, Sl., 2008. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. The Cochrane Collaboration. DOI:10.1002/9780470712184

Ho, T.C., Pham, H.T., Miller, J.G., Kircanski, K., Gotlib, I.H., 2020. Sympathetic nervous system dominance during stress recovery mediates associations between stress sensitivity and social anxiety symptoms in female adolescents. Dev. Psychopathol. 32, 1914–1925.

Hostinar, C.E., Johnson, A.E., Gunnar, M.R., 2015a. Parent support is less effective in buffering cortisol stress reactivity for adolescents compared to children. Dev. Sci. 18, 281–297. https://doi.org/10.1111/desc.12195

Hostinar, C.E., Johnson, A.E., Gunnar, M.R., 2015b. Early social deprivation and the social buffering of cortisol stress responses in late childhood: An experimental study. Dev. Psychol. 51, 1597–1608. https://doi.org/10.1037/dev0000029

Illius, S., Barthel, M.-C., Alexander, N., Vogel, S., 2024. The relevance of individual differences in coping styles for cortisol stress reactivity and habituation during repeated psychosocial stress [Poster presentation]. Presented at the Psychologie und Gehirn 2024, Hamburg.

Ji, J., Negriff, S., Kim, H., Susman, E.J., 2016. A study of cortisol reactivity and recovery among young adolescents: Heterogeneity and longitudinal stability and change. Dev. Psychobiol. 58, 283–302. https://doi.org/10.1002/dev.21369

Kassambara, A. (2023a). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.6.0, URL: https://CRAN.R-project.org/package=ggpubr .

Kassambara, A. (2023b). rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.2.999, URL: https://rpkgs.datanovia.com/rstatix/ .

Kelly, O., Matheson, K., Martinez, A., Merali, Z., Anisman, H., 2007. Psychosocial stress evoked by a virtual audience: relation to neuroendocrine activity. Cyberpsychol. Behav. 10, 655–662. https://doi.org/10.1089/cpb.2007.9973

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, 76–81. https://doi.org/10.1159/000119004

Kudielka, B.M., Buske-Kirschbaum, A., Hellhammer, D.H., Kirschbaum, C., 2004. Differential heart rate reactivity and recovery after psychosocial stress (TSST) in healthy children, younger adults, and elderly adults: The impact of age and gender. Int. J. Behav. Med. 11, 116–121. https://doi.org/10.1207/s15327558ijbm1102_8

Kudielka, B.M., Hellhammer, D.H., Kirschbaum, C., 2007. Ten Years of Research with the Trier Social Stress Test--Revisited, in: Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior. The Guilford Press, New York, NY, US, pp. 56–83.

Labuschagne, I., Grace, C., Rendell, P., Terrett, G., Heinrichs, M., 2019. An introductory guide to conducting the Trier Social Stress Test. Neurosci. Biobehav. Rev. 107, 686–695. https://doi.org/10.1016/j.neubiorev.2019.09.032

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

Lupien, S.J., King, S., Meaney, M.J., McEwen, B.S., 2000. Child's stress hormone levels correlate with mother's socioeconomic status and depressive state. Biol. Psychiatry 48, 976–980. https://doi.org/10.1016/s0006-3223(00)00965-3

Lüdecke, D. (2019). esc: Effect Size Computation for Meta Analysis (Version 0.5.1), URL: https://CRAN.R-project.org/package=esc .

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2023). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-13, URL: https://CRAN.R-project.org/package=e1071

Miller, R., Plessow, F., Kirschbaum, C., Stalder, T., 2013. Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: evaluation of salivary cortisol pulse detection in panel designs. Psychosomatic medicine, 75(9), 832-840.

Narvaez Linares, N.F., Charron, V., Ouimet, A.J., Labelle, P.R., Plamondon, H., 2020. A systematic review of the Trier Social Stress Test methodology: Issues in promoting study comparison and replicable research. Neurobiol. Stress 13, 100235. https://doi.org/10.1016/j.ynstr.2020.100235

Ordaz, S., Luna, B., 2012. Sex differences in physiological reactivity to acute psychosocial stress in adolescence. Psychoneuroendocrinology 37, 1135–1157. https://doi.org/10.1016/j.psyneuen.2012.01.002

Pennington, C. R. (2023). A student's guide to open science: using the replication crisis to reform psychology / Charlotte R. Pennington. Open University Press, McGraw Hill.

Pruessner, J. C., Wolf, O. T., Hellhammer, D. H., Buske-Kirschbaum, A., Von Auer, K., Jobst, S., ..., Kirschbaum, C., 1997. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. Life sciences, 61(26), 2539-2549.

Pustejovsky, J. (2023). clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections. R package version 0.5.10, URL: https://CRAN.R-project.org/package=clubSandwich

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R package version 2.3.6, URL: https://CRAN.R-project.org/package=psych .

Rohatgi, A., 2022. WebPlotDigitizer - Extract data from plots, images, and maps [WWW Document]. WebPlotDigitizer 46. URL https://automeris.io/WebPlotDigitizer/citation.html (accessed 9.6.23).

Rozenman, M., Vreeland, A., Iglesias, M., Mendez, M., Piacentini, J., 2018. The tell-tale heart: Physiological reactivity during resolution of ambiguity in youth anxiety. Cogn. Emot. 32, 389–396. https://doi.org/10.1080/02699931.2017.1289152

Schommer, N. C., Hellhammer, D. H., Kirschbaum, C., 2003. Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. Psychosomatic medicine, 65(3), 450-460.

Schwabe, L., Haddad, L., Schachinger, H., 2008. HPA axis activation by a socially evaluated cold-pressor test. Psychoneuroendocrinology, 33(6), 890-895.

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. https://doi.org/10.1016/j.psyneuen.2020.104582

Stalder, T., Lupien, S. J., Kudielka, B. M., Adam, E. K., Pruessner, J. C., Wüst, S., ..., Clow, A., 2022. Evaluation and update of the expert consensus guidelines for the assessment of the cortisol awakening response (CAR). Psychoneuroendocrinology, 146, 105946.

Stroud, L. R., Foster, E., Papandonatos, G. D., Handwerger, K., Granger, D. A., Kivlighan, K. T., Niaura, R., 2009. Stress response and the adolescent transition: Performance versus peer rejection stressors. Development and psychopathology, 21(1), 47-68.

Thabrew, H., Ruppeldt, P., Sollers, J.J., 2018. Systematic Review of Biofeedback Interventions for Addressing Anxiety and Depression in Children and Adolescents with Long-Term Physical Conditions. Appl. Psychophysiol. Biofeedback 43, 179–192. https://doi.org/10.1007/s10484-018-9399-z

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Journal of Statistical Software, 36(3), 1-48. https://doi.org/10.18637/jss.v036.i03

von Dawans, B., Kirschbaum, C., Heinrichs, M., 2011. The Trier Social Stress Test for Groups (TSST-G): A new research tool for controlled simultaneous social stress exposure in a group format. Psychoneuroendocrinology 36, 514–522. https://doi.org/10.1016/j.psyneuen.2010.08.004

Weber, J., Angerer, P., Apolinário-Hagen, J., 2022. Physiological reactions to acute stressors and subjective stress during daily life: A systematic review on ecological momentary assessment (EMA) studies. PLOS ONE 17, e0271996. https://doi.org/10.1371/journal.pone.0271996

Wickham, H. (2023). forcats: Tools for Working with Categorical Variables (Factors). R package version 1.0.0, URL: https://CRAN.R-project.org/package=forcats

Wüst, S., Federenko, I.S., van Rossum, E.F.C., Koper, J.W., Hellhammer, D.H., 2005. Habituation of cortisol responses to repeated psychosocial stress-further characterization and impact of genetic factors. Psychoneuroendocrinology 30, 199–211. https://doi.org/10.1016/j.psyneuen.2004.07.002

Yim, I.S., Quas, J.A., Rush, E.B., Granger, D.A., Skoluda, N., 2015. Experimental manipulation of the Trier Social Stress Test-Modified (TSST-M) to vary arousal across development. Psychoneuroendocrinology 57, 61–71. https://doi.org/10.1016/j.psyneuen.2015.03.021

Supplementary Online Materials

Experimental stress induction in children and adolescents with the Trier Social Stress Test (TSST): a systematic review and meta-analysis

Saskia Seel[1,3], Bernhard Pastötter[2,3], & Gregor Domes[1,3]

[1] Biological and Clinical Psychology, Department of Psychology, University of Trier, Trier, Germany

[2] Cognitive Psychology, Department of Psychology, University of Trier, Trier, Germany

[3] Institute for Cognitive and Affective Neuroscience, University of Trier, Trier, Germany

Corresponding author: Gregor Domes, University of Trier, Department of Biological and Clinical Psychology, Johanniterufer 15, 54290 Trier, Germany. *E-mail address*: domes@uni-trier.de; *Phone*: +49 651 201 2929; *Fax*: +49 651 201 2934

Full list of studies included in the systematic review and meta-analysis

[1] Studies included in the systematic review. [2] Studies included in the meta-analysis.

[1]Alen, N. V., Deer, L. K., Karimi, M., Feyzieva, E., Hastings, P. D., Hostinar, C. E., 2021. Children's altruism following acute stress: The role of autonomic nervous system activity and social support. Developmental Science, 24(5), e13099. https://doi.org/10.1111/desc.13099

[1,2]Alexander, N., Rosenloecher, F., Stalder, T., Linke, J., Distler, W., Morgner, J., Kirschbaum, C., 2012. Impact of antenatal synthetic glucocorticoid exposure on endocrine stress reactivity in term-born children. Journal of Clinical Endocrinology & Metabolism, 97(10), 3538–3544. https://doi.org/10.1210/jc.2012-1970

[1,2]Armbruster, D., Mueller, A., Strobel, A., Lesch, K.-P., Brocke, B., Kirschbaum, C., 2012. Children under stress—COMT genotype and stressful life events predict cortisol increase in an acute social stress paradigm. The International Journal of Neuropsychopharmacology, 15(9), 1229–1239. https://doi.org/10.1017/S1461145711001763

[1,2]Armbruster, D., Mueller-Alcazar, A., Strobel, A., Lesch, K.-P., Kirschbaum, C., Brocke, B., 2016. BDNF val(66)met genotype shows distinct associations with the acoustic startle reflex and the cortisol stress response in young adults and children. Psychoneuroendocrinology, 66, 39–46. https://doi.org/10.1016/j.psyneuen.2015.12.020

[1,2]Asbrand, J., Heinrichs, N., Nitschke, K., Wolf, O. T., Schmidtendorf, S., Tuschen-Caffier, B., 2019. Repeated stress leads to enhanced cortisol stress response in child social anxiety disorder but this effect can be prevented with CBT. Psychoneuroendocrinology, 109, 104352. https://doi.org/10.1016/j.psyneuen.2019.06.003

[1,2]Bae, Y. J., Gaudl, A., Jaeger, S., Stadelmann, S., Hiemisch, A., Kiess, W., Willenberg, A., Schaab, M., von Klitzing, K., Thiery, J., Ceglarek, U., Döhnert, M., Kratzsch, J., 2016. Immunoassay or LC-MS/MS for the measurement of salivary cortisol in children? Clinical Chemistry and Laboratory Medicine, 54(5), 811–822. https://doi.org/10.1515/cclm-2015-0412

[1,2]Başgöze, Z., Mirza, S. A., Silamongkol, T., Hill, D., Falke, C., Thai, M., Westlund Schreiner, M., Parenteau, A. M., Roediger, D. J., Hendrickson, T. J., Mueller, B. A., Fiecas, M. B., Klimes-Dougan, B., Cullen, K. R., 2021. Multimodal assessment of sustained threat in adolescents with nonsuicidal self-injury. Development and Psychopathology, 33(5), 1774–1792. https://doi.org/10.1017/S0954579421000754

[1]Bendezú, J. J., Calhoun, C. D., Wadsworth, M. E., 2022. Within-person patterns of psychobiological stress response correspondence: Links to preadolescent internalizing problems and coping behaviors. Anxiety, Stress, and Coping, 35(5), 592–608. https://doi.org/10.1080/10615806.2021.1982912

[1]Bendezu, J. J., Perzow, S. E. D., Wadsworth, M. E., 2016. What constitutes effective coping and efficient physiologic regulation following psychosocial stress depends on involuntary stress

responses. Psychoneuroendocrinology, 73, 42–50. https://doi.org/10.1016/j.psyneuen.2016.07.005

[1,2]Bendezú, J. J., Thai, M., Wiglesworth, A., Cullen, K. R., Klimes-Dougan, B., 2022. Adolescent stress experience-expression-physiology correspondence: Links to depression, self-injurious thoughts and behaviors, and frontolimbic neural circuity. Journal of Affective Disorders, 300, 269–279. https://doi.org/10.1016/j.jad.2021.12.098

[1]Bendezú, J. J., Wadsworth, M. E., 2018. Person-centered examination of salivary cortisol and alpha-amylase responses to psychosocial stress: Links to preadolescent behavioral functioning and coping. Biological Psychology, 132, 143–153. https://doi.org/10.1016/j.biopsycho.2017.11.011

[1]Bendezú, J. J., Wodzinski, A., Loughlin-Presnal, J. E., Mozeko, J., Cobler, S., Wadsworth, M. E., 2022. A multiple levels of analysis examination of the performance goal model of depression vulnerability in preadolescent children. Development and Psychopathology, 34(1), 241–261. https://doi.org/10.1017/S0954579420000851

[1,2]Bernhard, A., Ackermann, K., Martinelli, A., Chiocchetti, A. G., Vllasaliu, L., González-Madruga, K., Batchelor, M., Raschle, N. M., Oldenhof, H., Jansen, L. M. C., Kohls, G., Konrad, K., Popma, A., Stadler, C., Fairchild, G., Freitag, C. M., 2022. Neuroendocrine stress response in female and male youths with conduct disorder and associations with early adversity. Journal of the American Academy of Child and Adolescent Psychiatry, 61(5), 698–710. https://doi.org/10.1016/j.jaac.2021.11.023

[1,2]Bernhard, A., van der Merwe, C., Ackermann, K., Martinelli, A., Neumann, I. D., Freitag, C. M., 2018. Adolescent oxytocin response to stress and its behavioral and endocrine correlates. Hormones and Behavior, 105, 157–165. https://doi.org/10.1016/j.yhbeh.2018.08.010

[1,2]Bluth, K., Roberson, P. N. E., Gaylord, S. A., Faurot, K. R., Grewen, K. M., Arzon, S., Girdler, S. S., 2016. Does self-compassion protect adolescents from stress? Journal of Child and Family Studies, 25(4), 1098–1109. https://doi.org/10.1007/s10826-015-0307-3

[1]Brand, S., Mikoteit, T., Kalak, N., Sadeghi Bahmani, D., Lemola, S., Gerber, M., Ludyga, S., Bossard, M., Pühse, U., Holsboer-Trachsler, E., Hatzinger, M., 2018. Cortisol impacted on explicit learning encoding, but not on storage and retrieval, and was not associated with sleep patterns: Results from the Trier Social Stress Test for Children (TSST-C) among 9-years old children. Frontiers in Psychology, 9, 2240. https://doi.org/10.3389/fpsyg.2018.02240

[1,2]Buchweitz, A., de Azeredo, L., Esper, N., Dalfovo, N., Picoli, F., da Cunha, F., Viola, T., Grassi-Oliveira, R., 2023. Developmental dyslexia and the stress of reading: A social stress study of neuroendocrine response in children. MIND BRAIN AND EDUCATION. https://doi.org/10.1111/mbe.12361

[1,2]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, 419–426. https://doi.org/10.1097/00006842-199707000-00012

[1,2]Buske-Kirschbaum, A., Krieger, S., Wilkes, C., Rauh, W., Weiss, S., Hellhammer, D. H., 2007. Hypothalamic-pituitary-adrenal axis function and the cellular immune response in former preterm children. Journal of Clinical Endocrinology & Metabolism, 92(9), 3429–3435. https://doi.org/10.1210/jc.2006-2223

[1,2]Buske-Kirschbaum, A., Trikojat, K., Tesch, F., Schmitt, J., Roessner, V., Luksch, H., Rosen-Wolff, A., Plessow, F., 2019. Altered hypothalamus-pituitary-adrenal axis function: A relevant factor in the comorbidity of atopic eczema and attention deficit/hyperactivity disorder? Psychoneuroendocrinology, 105, 178–186. https://doi.org/10.1016/j.psyneuen.2018.12.005

[1,2]Buske-Kirschbaum, A., von Auer, K., Krieger, S., Weis, S., Rauh, W., Hellhammer, D., 2003. Blunted cortisol responses to psychosocial stress in asthmatic children: A general feature of atopic disease? Psychosomatic Medicine, 65(5), 806. https://doi.org/10.1097/01.PSY.0000095916.25975.4F

[1,2]Busso, D. S., McLaughlin, K. A., Sheridan, M. A., 2017. Dimensions of adversity, physiological reactivity, and externalizing psychopathology in adolescence: Deprivation and threat. Psychosomatic Medicine, 79(2), 162–171. https://doi.org/10.1097/PSY.0000000000000369

[1,2]Cameron, C.A., McKay, S., Susman, E.J., Wynne-Edwards, K., Wright, J.M., Weinberg, J., 2017. Cortisol Stress Response Variability in Early Adolescence: Attachment, Affect and Sex. J. Youth Adolesc. 46, 104–120. https://doi.org/10.1007/s10964-016-0548-5

[1,2]Chaplin, T. M., Freiburger, M. B., Mayes, L. C., Sinha, R., 2010. Prenatal cocaine exposure, gender, and adolescent stress response: A prospective longitudinal study. Neurotoxicology and Teratology, 32(6), 595–604. https://doi.org/10.1016/j.ntt.2010.08.007

[1,2]Chen, F. R., Raine, A., Granger, D. A., 2020. The within-person coordination of HPA and ANS activity in stress response: Relation with behavior problems. Psychoneuroendocrinology, 121, 104805. https://doi.org/10.1016/j.psyneuen.2020.104805

[1,2]Chen, F. S., Schmitz, J., Domes, G., Tuschen-Caffier, B., Heinrichs, M., 2014. Effects of acute social stress on emotion processing in children. Psychoneuroendocrinology, 40, 91–95. https://doi.org/10.1016/j.psyneuen.2013.11.003

[1]Chen, G., Kong, Y., Deater-Deckard, K., Zhang, W., 2018. Bullying victimization heightens cortisol response to psychosocial stress in Chinese children. Journal of Abnormal Child Psychology, 46(5), 1051–1059. https://doi.org/10.1007/s10802-017-0366-6

[1,2]Chen, L., Zilioli, S., Jiang, Y., Wang, X., Lin, D., 2021. Perceived social support and children's physiological responses to stress: An examination of the stress-buffering hypothesis. Psychosomatic Medicine, 83(1), 51–61. https://doi.org/10.1097/PSY.0000000000000875

[1]Chubar, V., Vaessen, T., Noortgate, W. V. den, Lutin, E., Bosmans, G., Bekaert, B., Van Leeuwen, K., Calders, F., Weyn, S., Bijttebier, P., Goossens, L., Claes, S., 2023. Mild daily stress, in interaction with NR3C1 DNA methylation levels, is linked to alterations in the HPA axis and ANS response to acute stress in early adolescents. Psychoneuroendocrinology, 150, 106045. https://doi.org/10.1016/j.psyneuen.2023.106045

[1]Cook, E. C., Chaplin, T. M., Sinha, R., Tebes, J. K., Mayes, L. C., 2012. The stress response and adolescents' adjustment: The impact of child maltreatment. Journal of Youth and Adolescence, 41(8), 1067–1077. https://doi.org/10.1007/s10964-012-9746-y

[1,2]Corbett, B. A., Muscatello, R. A., Kim, A., Patel, K., Vandekar, S., 2021. Developmental effects in physiological stress in early adolescents with and without autism spectrum disorder. Psychoneuroendocrinology, 125, 105115. https://doi.org/10.1016/j.psyneuen.2020.105115

[1]Crossman, M. K., Kazdin, A. E., Kitt, E. R., 2018. The influence of a socially assistive robot on mood, anxiety, and arousal in children. Professional Psychology-Research and Practice, 49(1), 48–56. https://doi.org/10.1037/pro0000177

[1,2]Crossman, M. K., Kazdin, A. E., Matijczak, A., Kitt, E. R., Santos, L. R., 2020. The influence of interactions with dogs on affect, anxiety, and arousal in children. Journal of Clinical Child and Adolescent Psychology: 49(4), 535–548. https://doi.org/10.1080/15374416.2018.1520119

[1,2]Dockray, S., Susman, E. J., Dorn, L. D., 2009. Depression, cortisol reactivity, and obesity in childhood and adolescence. Journal of Adolescent Health, 45(4), 344–350. https://doi.org/10.1016/j.jadohealth.2009.06.014

[1]Dorn, L. D., Campo, J. C., Thato, S., Dahl, R. E., Lewin, D., Chandra, R., Di Lorenzo, C., 2003. Psychological comorbidity and stress reactivity in children and adolescents with recurrent abdominal pain and anxiety disoders. Journal of the American Academy of Child & Adolescent Psychiatry, 42(1), 66–75. https://doi.org/10.1097/00004583-200301000-00012

[1,2]Edmiston, E. K., Blain, S. D., Corbett, B. A., 2017. Salivary cortisol and behavioral response to social evaluative threat in adolescents with autism spectrum disorder. Autism Research, 10(2), 346–358. https://doi.org/10.1002/aur.1660

[1,2]Erni, K., Shaqiri-Emini, L., La Marca, R., Zimmermann, R., Ehlert, U., 2012. Psychobiological effects of prenatal glucocorticoid exposure in 10-year-old-children. Frontiers in Psychiatry, 3, 104. https://doi.org/10.3389/fpsyt.2012.00104

[1,2]Espin, L., García, I., Del Pino Sánchez, M., Román, F., Salvador, A., 2019. Effects of psychosocial stress on the hormonal and affective response in children with dyslexia. Trends in Neuroscience and Education, 15, 1–9. https://doi.org/10.1016/j.tine.2019.03.001

[1,2]Ewing-Cobbs, L., Prasad, M. R., Cox, C. S., Granger, D. A., Duque, G., Swank, P. R., 2017. Altered stress system reactivity after pediatric injury: Relation with post traumatic stress symptoms. Psychoneuroendocrinology, 84, 66–75. https://doi.org/10.1016/j.psyneuen.2017.06.003

[1,2]Fearon, R. M. P., Tomlinson, M., Kumsta, R., Skeen, S., Murray, L., Cooper, P. J., Morgan, B., 2017. Poverty, early care, and stress reactivity in adolescence: Findings from a prospective, longitudinal study in South Africa. Development and Psychopathology, 29(2), 449–464. https://doi.org/10.1017/S0954579417000104

[1]Fisher, P. A., Kim, H. K., Bruce, J., Pears, K. C., 2012. Cumulative effects of prenatal substance exposure and early adversity on foster children's HPA-axis reactivity during a psychosocial stressor. International Journal of Behavioral Development, 36(1), 29–35. https://doi.org/10.1177/0165025411406863

[1]Fu, J., Wang, H., Weng, T., Tao, F., 2012. Cortisol response to the Trier Social Stress Test among Chinese adolescents. Journal of Pediatric Endocrinology & Metabolism, 25(11–12), 1213–1216. https://doi.org/10.1515/jpem-2012-0234

[1,2]Funke, R., Eichler, A., Distler, J., Golub, Y., Kratz, O., Moll, G. H., 2017. Stress system dysregulation in pediatric generalized anxiety disorder associated with comorbid depression. Stress and Health: Journal of the International Society for the Investigation of Stress, 33(5), 518–529. https://doi.org/10.1002/smi.2736

[1,2]Geronikolou, S. A., Chamakou, A., Mantzou, A., Chrousos, G., Kanaka-Gantenbein, C., 2015. Frequent cellular phone use modifies hypothalamic-pituitary-adrenal axis response to a cellular phone call after mental stress in healthy children and adolescents: A pilot study. Science of the Total Environment, 536, 182–188. https://doi.org/10.1016/j.scitotenv.2015.07.052

[1,2]Geronikolou, S. A., Vasdekis, V., Mantzou, A., Davos, C., Cokkinos, D. V., Chrousos, G. P., 2023. Neuroendocrine system adaptation during consecutive extrinsic stimuli: A pilot dynamic study. Children (Basel, Switzerland), 10(2). https://doi.org/10.3390/children10020248

[1,2]Glier, S., Campbell, A., Corr, R., Pelletier-Baldelli, A., Yefimov, M., Guerra, C., Scott, K., Murphy, L., Bizzell, J., Belger, A., 2022. Coordination of autonomic and endocrine stress responses to the Trier Social Stress Test in adolescence. Psychophysiology, 59(9), e14056. https://doi.org/10.1111/psyp.14056

[1,2]Gordijn, M. S., Gemke, R. J. B. J., Bierings, M. B., Hoogerbrugge, P. M., Tersteeg-Kamperman, M. D. J., Heijnen, C. J., Rotteveel, J., Kaspers, G. J. L., 2013. Adequate endocrine and cardiovascular response to social stress in survivors of childhood acute lymphoblastic leukemia. Psychoneuroendocrinology, 38(12), 3145–3149. https://doi.org/10.1016/j.psyneuen.2013.07.013

[1]Gordis, E.B., Granger, D.A., Susman, E.J., Trickett, P.K., 2006. Asymmetry between salivary cortisol and alpha-amylase reactivity to stress: relation to aggressive behavior in adolescents. Psychoneuroendocrinology 31, 976–987. https://doi.org/10.1016/j.psyneuen.2006.05.010

[1,2]Gulewitsch, M. D., Weimer, K., Enck, P., Schwille-Kiuntke, J., Hautzinger, M., Schlarb, A. A., 2017. Stress reactivity in childhood functional abdominal pain or irritable bowel syndrome. European Journal of Pain, 21(1), 166–177. https://doi.org/10.1002/ejp.914

[1,2]Gunnar, M. R., Frenn, K., Wewerka, S. S., Van Ryzin, M. J., 2009. Moderate versus severe early life stress: Associations with stress reactivity and regulation in 10-12-year-old children. Psychoneuroendocrinology, 34(1), 62–75. https://doi.org/10.1016/j.psyneuen.2008.08.013

[1,2]Gunnar, M.R., Reid, B.M., Donzella, B., Miller, Z.R., Gardow, S., Tsakonas, N.C., Thomas, K.M., DeJoseph, M., Bendezú, J.J., 2021. Validation of an online version of the Trier Social Stress Test in a study of adolescents. Psychoneuroendocrinology 125, 105111. https://doi.org/10.1016/j.psyneuen.2020.105111

[1,2]Gunnar, M. R., Wewerka, S., Frenn, K., Long, J. D., Griggs, C., 2009. Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: Normative changes and associations with puberty. Development and Psychopathology, 21(1), 69–85. https://doi.org/10.1017/S0954579409000054

[1]Hackman, D. A., Betancourt, L. M., Brodsky, N. L., Hurt, H., Farah, M. J., 2012. Neighborhood disadvantage and adolescent stress reactivity. Frontiers in Human Neuroscience, 6, 277. https://doi.org/10.3389/fnhum.2012.00277

[1,2]Hanke, M., Looser, V. N., Bruggisser, F., Leuenberger, R., Gerber, M., Ludyga, S., 2023. Moderate-to-vigorous physical activity and reactivity to acute psychosocial stress in preadolescent children. Journal of Science and Medicine in Sport, S1440-2440(23)00182-2. https://doi.org/10.1016/j.jsams.2023.07.010

[1,2]Hipwell, A. E., Keenan, K., Marsland, A., 2009. Exploring psychophysiological markers of vulnerability to somatic illnesses in females. Journal of Pediatric Psychology, 34(9), 1030–1039. https://doi.org/10.1093/jpepsy/jsp010

[1,2]Houbrechts, M., Cuyvers, B., Goossens, L., Bijttebier, P., Bröhl, A. S., Calders, F., Chubar, V., Claes, S., Geukens, F., Van Leeuwen, K., Noortgate, W. V. D., Weyn, S., Bosmans, G., 2023. Parental support and insecure attachment development: The cortisol stress response as a moderator. Attachment & Human Development, 25(1), 104–116. https://doi.org/10.1080/14616734.2021.1907968

[1]Howland, M. A., Donzella, B., Miller, B. S., Gunnar, M. R., 2020. Pubertal recalibration of cortisol-DHEA coupling in previously-institutionalized children. Hormones and Behavior, 125, 104816. https://doi.org/10.1016/j.yhbeh.2020.104816

[1,2]Joos, C. M., McDonald, A., Wadsworth, M. E., 2019. Extending the toxic stress model into adolescence: Profiles of cortisol reactivity. Psychoneuroendocrinology, 107, 46–58. https://doi.org/10.1016/j.psyneuen.2019.05.002

[1,2]Kaess, M., Hille, M., Parzer, P., Maser-Gluth, C., Resch, F., Brunner, R., 2012. Alterations in the neuroendocrinological stress response to acute psychosocial stress in adolescents

engaging in nonsuicidal self-injury. Psychoneuroendocrinology, 37(1), 157–161. https://doi.org/10.1016/j.psyneuen.2011.05.009

[1,2]Kater, M.-J., Werner, A., Schlarb, A. A., Lohaus, A., 2022. Stress reactivity in salivary cortisol and electrocardiogram in adolescents: Investigating sleep disturbances and insomnia. Journal of Sleep Research, 31(6), e13591. https://doi.org/10.1111/jsr.13591

[1]Kertes, D. A., Liu, J., Hall, N. J., Hadad, N. A., Wynne, C. D. L., Bhatt, S. S., 2017. Effect of pet dogs on children's perceived stress and cortisol stress response. Social Development, 26(2), 382–401. https://doi.org/10.1111/sode.12203

[1,2]Kidwell, K. M., Reiter-Purtill, J., Decker, K., Howarth, T., Doland, F., Zeller, M. H., 2022. Stress and eating responses in adolescent females predisposed to obesity: A pilot and feasibility study. Appetite, 179, 106308. https://doi.org/10.1016/j.appet.2022.106308

[1]Kircanski, K., Sisk, L. M., Ho, T. C., Humphreys, K. L., King, L. S., Colich, N. L., Ordaz, S. J., Gotlib, I. H., 2019. Early life stress, cortisol, frontolimbic connectivity, and depressive symptoms during puberty. Development and Psychopathology, 31(3), 1011–1022. https://doi.org/10.1017/S0954579419000555

[1,2]Koenig, J., Lischke, A., Bardtke, K., Heinze, A.-L., Kröller, F., Pahnke, R., Kaess, M., 2022. Altered psychobiological reactivity but no impairment of emotion recognition following stress in adolescents with non-suicidal self-injury. European Archives of Psychiatry and Clinical Neuroscience. https://doi.org/10.1007/s00406-022-01496-4

[1,2]Kraemer, M., Seefeldt, W. L., Heinrichs, N., Tuschen-Caffier, B., Schmitz, J., Wolf, O. T., Blechert, J., 2012. Subjective, autonomic, and endocrine reactivity during social stress in children with social phobia. Journal of Abnormal Child Psychology, 40(1), 95–104. https://doi.org/10.1007/s10802-011-9548-9

[1]Krishnaveni, G. V., Jones, A., Veena, S. R., Somashekara, R., Karat, S. C., Fall, C. H. D., 2018. Adiposity and cortisol response to stress in Indian adolescents. Indian Pediatrics, 55(2), 125–130. https://doi.org/10.1007/s13312-018-1244-1

[1]Krishnaveni, G. V., Veena, S. R., Jones, A., Bhat, D. S., Malathi, M. P., Hellhammer, D., Srinivasan, K., Upadya, H., Kurpad, A. V., Fall, C. H. D., 2014. Trier social stress test in Indian adolescents. Indian Pediatrics, 51(6), 463–467. https://doi.org/10.1007/s13312-014-0437-5

[1]Krishnaveni, G. V., Veena, S. R., Jones, A., Srinivasan, K., Osmond, C., Karat, S. C., Kurpad, A. V., Fall, C. H. D., 2015. Exposure to maternal gestational diabetes is associated with higher cardiovascular responses to stress in adolescent Indians. Journal of Clinical Endocrinology & Metabolism, 100(3), 986–993. https://doi.org/10.1210/jc.2014-3239

[1]Krishnaveni, G., Veena, S. R., Johnson, M., Kumaran, K., Jones, A., Bhat, D. S., Yajnik, C. S., Fall, C. H. D., 2020. Maternal B12, folate and homocysteine concentrations and offspring cortisol and cardiovascular responses to stress. Journal of Clinical Endocrinology & Metabolism, 105(7), E2591–E2599. https://doi.org/10.1210/clinem/dgz114

[1,2]Kuhlman, K. R., Repetti, R. L., Reynolds, B. M., Robles, T. F., 2018. Interparental conflict and child HPA-axis responses to acute stress: Insights using intensive repeated measures. Journal of Family Psychology, 32(6), 773–782. https://doi.org/10.1037/fam0000437

[1]Lanni, K. E., Schupp, C. W., Simon, D., Corbett, B. A., 2012. Verbal ability, social stress, and anxiety in children with autistic disorder. Autism: The International Journal of Research and Practice, 16(2), 123–138. https://doi.org/10.1177/1362361311425916

[1]Laurent, H. K., Gilliam, K. S., Wright, D. B., Fisher, P. A., 2015. Child anxiety symptoms related to longitudinal cortisol trajectories and acute stress responses: Evidence of developmental stress sensitization. Journal of Abnormal Psychology, 124(1), 68–79. https://doi.org/10.1037/abn0000009

[1,2]Liu, J., Chaplin, T. M., Wang, F., Sinha, R., Mayes, L. C., Blumberg, H. P., 2012. Stress reactivity and corticolimbic response to emotional faces in adolescents. Journal of the American Academy of Child and Adolescent Psychiatry, 51(3), 304–312. https://doi.org/10.1016/j.jaac.2011.12.014

[1]Lu, Q., Pan, F., Ren, L., Xiao, J., Tao, F., 2018. Sex differences in the association between internalizing symptoms and hair cortisol level among 10–12-year-old adolescents in China. PloS One, 13(3), e0192901. https://doi.org/10.1371/journal.pone.0192901

[1]Lu, Q., Tao, F., Hou, F., Zhang, Z., Sun, Y., Xu, Y., Xu, S., Zhao, Y., 2014. Cortisol reactivity, delay discounting and percent body fat in Chinese urban young adolescents. Appetite, 72, 13–20. https://doi.org/10.1016/j.appet.2013.09.019

[1,2]Lucas-Thompson, R. G., Granger, D. A., 2014. Parent-child relationship quality moderates the link between marital conflict and adolescents' physiological responses to social evaluative threat. Journal of Family Psychology, 28(4), 538–548. https://doi.org/10.1037/a0037328

[1,2]Mackrell, S. V. M., Sheikh, H. I., Kotelnikova, Y., Kryski, K. R., Jordan, P. L., Singh, S. M., Hayden, E. P., 2014. Child temperament and parental depression predict cortisol reactivity to stress in middle childhood. Journal of Abnormal Psychology, 123(1), 106–116. https://doi.org/10.1037/a0035612

[1]MacMillan, H. L., Georgiades, K., Duku, E. K., Shea, A., Steiner, M., Niec, A., Tanaka, M., Gensey, S., Spree, S., Vella, E., Walsh, C. A., De Bellis, M. D., Van der Meulen, J., Boyle, M. H., Schmidt, L. A., 2009. Cortisol response to stress in female youths exposed to childhood maltreatment: Results of the youth mood project. Biological Psychiatry, 66(1), 62–68. https://doi.org/10.1016/j.biopsych.2008.12.014

[1,2]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. Molecular Psychiatry, 26(9), 4823–4838. https://doi.org/10.1038/s41380-020-0747-z

[1,2]Marques-Feixa, L., Palma-Gudiel, H., Romero, S., Moya-Higueras, J., Rapado-Castro, M., Castro-Quintas, Á., Zorrilla, I., José Muñoz, M., Ramírez, M., Mayoral, M., Mas, A., José Lobato, M., Blasco-Fontecilla, H., Fañanás, L., EPI-Young Stress GROUP, 2021. Childhood maltreatment disrupts HPA-axis activity under basal and stress conditions in a dose-response relationship in children and adolescents. Psychological Medicine, 1–14. https://doi.org/10.1017/S003329172100249X

[1]Martel, F. L., Hayward, C., Lyons, D. M., Sanborn, K., Varady, S., Schatzberg, A. F., 1999. Salivary cortisol levels in socially phobic adolescent girls. Depression and Anxiety, 10(1), 25–27. https://doi.org/10.1002/(sici)1520-6394(1999)10:1<25::aid-da4>3.0.co;2-o

[1,2]Martikainen, S., Pesonen, A.-K., Lahti, J., Heinonen, K., Feldt, K., Pyhälä, R., Tammelin, T., Kajantie, E., Eriksson, J. G., Strandberg, T. E., Räikkönen, K., 2013. Higher levels of physical activity are associated with lower hypothalamic-pituitary-adrenocortical axis reactivity to psychosocial stress in children. The Journal of Clinical Endocrinology and Metabolism, 98(4), E619-627. https://doi.org/10.1210/jc.2012-3745

[1,2]Martin, A., Hellhammer, J., Hero, T., Max, H., Schult, J., Terstegen, L., 2011. Effective prevention of stress-induced sweating and axillary malodour formation in teenagers. International Journal of Cosmetic Science, 33(1), 90–97. https://doi.org/10.1111/j.1468-2494.2010.00596.x

[1]Martinez-Torteya, C., Bogat, G. A., Levendosky, A. A., von Eye, A., 2016. The influence of prenatal intimate partner violence exposure on hypothalamic-pituitary-adrenal axis reactivity and childhood internalizing and externalizing symptoms. Development and Psychopathology, 28(1), 55–72. https://doi.org/10.1017/S0954579415000280

[1,2]Mazurka, R., Wynne-Edwards, K. E., Harkness, K. L., 2018. Sex differences in the cortisol response to the Trier Social Stress Test in depressed and nondepressed adolescents. Clinical Psychological Science, 6(3), 301–314. https://doi.org/10.1177/2167702617739973

[1,2]McKay, S. L., Fouladirad, S., Cameron, C. A., 2021. The Frustration Social Stressor for Adolescents (FSS-A): A newly adapted psychosocial stressor. Stress and Health, 37(4), 715–728. https://doi.org/10.1002/smi.3029

[1,2]Morris, M. C., Kouros, C. D., Mielock, A. S., Rao, U., 2017. Depressive symptom composites associated with cortisol stress reactivity in adolescents. Journal of Affective Disorders, 210, 181–188. https://doi.org/10.1016/j.jad.2016.12.023

[1,2]Moss, R. H., Conner, M., O'Connor, D. B., 2020. Exploring the effects of daily hassles on eating behaviour in children: The role of cortisol reactivity. Psychoneuroendocrinology, 117, 104692. https://doi.org/10.1016/j.psyneuen.2020.104692

[1,2]Mrug, S., Tyson, A., Turan, B., Granger, D. A., 2016. Sleep problems predict cortisol reactivity to stress in urban adolescents. Physiology & Behavior, 155, 95–101. https://doi.org/10.1016/j.physbeh.2015.12.003

[1,2]Mueller, A., Armbruster, D., Moser, D. A., Canli, T., Lesch, K.-P., Brocke, B., Kirschbaum, C., 2011. Interaction of serotonin transporter gene-linked polymorphic region and stressful life events predicts cortisol stress response. Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 36(7), 1332–1339. https://doi.org/10.1038/npp.2011.11

[1,2]Mueller, S., Moser, D., Frach, L., Wimberger, P., Nitzsche, K., Li, S.-C., Kirschbaum, C., Alexander, N., 2022. No long-term effects of antenatal synthetic glucocorticoid exposure on epigenetic regulation of stress-related genes. Translational Psychiatry, 12(1), 62. https://doi.org/10.1038/s41398-022-01828-x

[1,2]Parenteau, A. M., Alen, N. V., Deer, L. K., Nissen, A. T., Luck, A. T., Hostinar, C. E., 2020. Parenting matters: Parents can reduce or amplify children's anxiety and cortisol responses to acute stress. Development and Psychopathology, 32(5), 1799–1809. https://doi.org/10.1017/S0954579420001285

[1,2]Peckins, M. K., Susman, E. J., Negriff, S., Noll, J., Trickett, P. K., 2015. Cortisol profiles: A test for adaptive calibration of the stress response system in maltreated and nonmaltreated youth. Development and Psychopathology, 27(4 Pt 2), 1461–1470. https://doi.org/10.1017/S0954579415000875

[1,2]Perry, N. B., Donzella, B., Parenteau, A. M., Desjardins, C., Gunnar, M. R., 2019. Emotion regulation and cortisol reactivity during a social evaluative stressor: A study of post-institutionalized youth. Developmental Psychobiology, 61(4), 557–572. https://doi.org/10.1002/dev.21828

[1]Pesonen, A.-K., Kajantie, E., Heinonen, K., Pyhala, R., Lahti, J., Jones, A., Matthews, K. A., Eriksson, J. G., Strandberg, T., Raikkonen, K., 2012. Sex-specific associations between sleep problems and hypothalamic-pituitary-adrenocortical axis activity in children. Psychoneuroendocrinology, 37(2), 238–248. https://doi.org/10.1016/j.psyneuen.2011.06.008

[1,2]Pham, H. T., Bendezú, J. J., Wadsworth, M. E., 2023. HPA-SAM co-activation among racially diverse, economically disadvantaged early adolescents: Secondary analysis with a preliminary test of a multisystem, person-centered approach. Biological Psychology, 179, 108546. https://doi.org/10.1016/j.biopsycho.2023.108546

[1,2]Ping, E. Y., Laplante, D. P., Elgbeili, G., Jones, S. L., Brunet, A., King, S., 2020. Disaster-related prenatal maternal stress predicts HPA reactivity and psychopathology in adolescent offspring: Project Ice Storm. Psychoneuroendocrinology, 117, 104697. https://doi.org/10.1016/j.psyneuen.2020.104697

[1,2]Pollak, K. M., Shao, S., Knutson, J. M., Gill, H., Lee, Y.-V., McKay, S. L., Cameron, C. A., 2022. Adolescent psychological and physiological responses to frustration- and anxiety-provoking stressors. Current Psychology, 41(1), 395–405. https://doi.org/10.1007/s12144-019-00582-6

[1]Portnoy, J., Raine, A., Glenn, A. L., Chen, F. R., Choy, O., Granger, D. A., 2015. Digit ratio (2d:4d) moderates the relationship between cortisol reactivity and self-reported externalizing behavior in young adolescent males. Biological Psychology, 112, 94–106. https://doi.org/10.1016/j.biopsycho.2015.09.013

[1]Quas, J. A., Rush, E. B., Yim, I. S., Edelstein, R. S., Otgaar, H., Smeets, T., 2016. Stress and emotional valence effects on children's versus adolescents' true and false memory. Memory, 24(5), 696–707. https://doi.org/10.1080/09658211.2015.1045909

[1]Quas, J. A., Rush, E. B., Yim, I. S., Nikolayev, M., 2014. Effects of stress on memory in children and adolescents: Testing causal connections. Memory, 22(6), 616–632. https://doi.org/10.1080/09658211.2013.809766

[1]Quas, J. A., Yim, I. S., Edelstein, R. S., Cahill, L., Rush, E. B., 2011. The role of cortisol reactivity in children's and adults' memory of a prior stressful experience. Developmental Psychobiology, 53(2), 166–174. https://doi.org/10.1002/dev.20505

[1]Quas, J. A., Yim, I. S., Rush, E., Sumaroka, M., 2012. Hypothalamic pituitary adrenal axis and sympathetic activation: Joint predictors of memory in children, adolescents, and adults. Biological Psychology, 89(2), 335–341. https://doi.org/10.1016/j.biopsycho.2011.11.006

[1,2]Quesada, A. A., Tristao, R. M., Pratesi, R., Wolf, O. T., 2014. Hyper-responsiveness to acute stress, emotional problems and poorer memory in former preterm children. Stress: The International Journal on the Biology of Stress, 17(5), 389–399. https://doi.org/10.3109/10253890.2014.949667

[1,2]Quesada, A. A., Wiemers, U. S., Schoofs, D., Wolf, O. T., 2012. Psychosocial stress exposure impairs memory retrieval in children. Psychoneuroendocrinology, 37(1), 125–136. https://doi.org/10.1016/j.psyneuen.2011.05.013

[1,2]Raffington, L., Falck, J., Heim, C., Mather, M., Shing, Y. L., 2020. Effects of stress on 6- and 7-year-old children's emotional memory differs by gender. Journal of Experimental Child Psychology, 199, 104924. https://doi.org/10.1016/j.jecp.2020.104924

[1,2]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. Developmental Psychology, 58(10), 1832–1848. https://doi.org/10.1037/dev0001393

[1]Raikkonen, K., Matthews, K. A., Pesonen, A.-K., Pyhala, R., Paavonen, E. J., Feldt, K., Jones, A., Phillips, D. I. W., Seckl, J. R., Heinonen, K., Lahti, J., Komsi, N., Jarvenpaa, A.-L., Eriksson, J. G., Strandberg, T. E., Kajantie, E., 2010. Poor sleep and altered Hypothalamic-Pituitary-Adrenocortical and Sympatho-Adrenal-Medullary System activity in children. Journal of Clinical Endocrinology & Metabolism, 95(5), 2254–2261. https://doi.org/10.1210/jc.2009-0943

[1]Räikkönen, K., Seckl, J. R., Heinonen, K., Pyhälä, R., Feldt, K., Jones, A., Pesonen, A.-K., Phillips, D. I. W., Lahti, J., Järvenpää, A.-L., Eriksson, J. G., Matthews, K. A., Strandberg,

T. E., Kajantie, E., 2010. Maternal prenatal licorice consumption alters hypothalamic-pituitary-adrenocortical axis function in children. Psychoneuroendocrinology, 35(10), 1587–1593. https://doi.org/10.1016/j.psyneuen.2010.04.010

[1]Rakers, F., Schleußner, E., Muth, I., Hoyer, D., Rupprecht, S., Schiecke, K., Groten, T., Dreiling, M., Kozik, V., Schwab, M., Hoyer, H., Ligges, C., 2022. Association between antenatal glucocorticoid exposure and the activity of the stress system, cognition, and behavior in 8- to 9-year-old children: A prospective observational study. Acta Obstetricia Et Gynecologica Scandinavica, 101(9), 996–1006. https://doi.org/10.1111/aogs.14386

[1]Randazzo, W. T., Dockray, S., Susman, E. J., 2008. The stress response in adolescents with inattentive type ADHD symptoms. Child Psychiatry and Human Development, 39(1), 27–38. https://doi.org/10.1007/s10578-007-0068-3

[1,2]Rao, U., Hammen, C., Ortiz, L. R., Chen, L.-A., Poland, R. E., 2008. Effects of early and recent adverse experiences on adrenal response to psychosocial stress in depressed adolescents. Biological Psychiatry, 64(6), 521–526. https://doi.org/10.1016/j.biopsych.2008.05.012

[1,2]Raymond, C., Pichette, F., Beaudin, M., Cernik, R., Marin, M.-F., 2023. Vulnerability to anxiety differently predicts cortisol reactivity and state anxiety during a laboratory stressor in healthy girls and boys. Journal of Affective Disorders, 331, 425–433. https://doi.org/10.1016/j.jad.2023.02.154

[1]Ruttle, P. L., Klein, M. H., Slattery, M. J., Kalin, N. H., Armstrong, J. M., Essex, M. J., 2014. Adolescent adrenocortical activity and adiposity: Differences by sex and exposure to early maternal depression. Psychoneuroendocrinology, 47, 68–77. https://doi.org/10.1016/j.psyneuen.2014.04.025

[1]Sato, A. F., Darling, K. E., Ruzicka, E. B., Fahrenkamp, A. J., Shields, C. V., West, C. E., Schuler, S. M., 2023. A biobehavioral examination of emotional eating in adolescents: Evaluating the role of acute cortisol stress reactivity in the context of a multi-method assessment of eating in response to stress. Appetite, 106486. https://doi.org/10.1016/j.appet.2023.106486

[1,2]Seltzer, L. J., Ziegler, T., Connolly, M. J., Prososki, A. R., Pollak, S. D., 2014. Stress-induced elevation of oxytocin in maltreated children: Evolution, neurodevelopment, and social behavior. Child Development, 85(2), 501–512. https://doi.org/10.1111/cdev.12136

[1,2]Shapero, B. G., McClung, G., Bangasser, D. A., Abramson, L. Y., Alloy, L. B., 2017. Interaction of biological stress recovery and cognitive vulnerability for depression in adolescence. Journal of Youth and Adolescence, 46(1), 91–103. https://doi.org/10.1007/s10964-016-0451-0

[1,2]Sheridan, M. A., How, J., Araujo, M., Schamberg, M. A., Nelson, C. A., 2013. What are the links between maternal social status, hippocampal function, and HPA axis function in children? Developmental Science, 16(5), 665–675. https://doi.org/10.1111/desc.12087

[1]Stadelmann, S., Jaeger, S., Matuschek, T., Bae, Y. J., von Klitzing, K., Klein, A. M., Döhnert, M., 2018. Endocrinological and subjective stress responses in children with depressive, anxiety,

or externalizing disorders. Development and Psychopathology, 30(2), 605–622. https://doi.org/10.1017/S0954579417001146

[1,2]Stewart, J. G., Mazurka, R., Bond, L., Wynne-Edwards, K. E., Harkness, K. L., 2013. Rumination and impaired cortisol recovery following a social stressor in adolescent depression. Journal of Abnormal Child Psychology, 41(7), 1015–1026. https://doi.org/10.1007/s10802-013-9740-1

[1,2]Strahler, J., Mueller, A., Rosenloecher, F., Kirschbaum, C., Rohleder, N., 2010. Salivary alpha-amylase stress reactivity across different age groups. Psychophysiology, 47(3), 587–595. https://doi.org/10.1111/j.1469-8986.2009.00957.x

[1,2]Sumner, J. A., McLaughlin, K. A., Walsh, K., Sheridan, M. A., Koenen, K. C., 2014. CRHR1 genotype and history of maltreatment predict cortisol reactivity to stress in adolescents. Psychoneuroendocrinology, 43, 71–80. https://doi.org/10.1016/j.psyneuen.2014.02.002

[1]Susman, E. J., Dockray, S., Granger, D. A., Blades, K. T., Randazzo, W., Heaton, J. A., Dorn, L. D., 2010. Cortisol and alpha amylase reactivity and timing of puberty: Vulnerabilities for antisocial behaviour in young adolescents. Psychoneuroendocrinology, 35(4), 557–569. https://doi.org/10.1016/j.psyneuen.2009.09.004

[1,2]Tackett, J. L., Herzhoff, K., Smack, A. J., Reardon, K. W., Adam, E. K., 2017. Does socioeconomic status mediate racial differences in the cortisol response in middle childhood? Health Psychology, 36(7), 662–672. https://doi.org/10.1037/hea0000480

[1]Tackett, J. L., Kushner, S. C., Josephs, R. A., Harden, K. P., Page-Gould, E., Tucker-Drob, E. M., 2014. Hormones: Empirical contribution: Cortisol reactivity and recovery in the context of adolescent personality disorder. Special Issue: A Developmental Psychopathology Perspective on Personality Disorder, 28(1), 25–39. https://doi.org/10.1521/pedi.2014.28.1.25

[1]Tarbell, S. E., Millar, A., Laudenslager, M., Palmer, C., Fortunato, J. E., 2017. Anxiety and physiological responses to the Trier Social Stress Test for Children in adolescents with cyclic vomiting syndrome. Autonomic Neuroscience: Basic & Clinical, 202, 79–85. https://doi.org/10.1016/j.autneu.2016.08.010

[1]Taylor-Cavelier, S. J., Micol, V. J., Roberts, A. G., Geiss, E. G., Lopez-Duran, N., 2021. DHEA moderates the impact of childhood trauma on the HPA Axis in adolescence. Neuropsychobiology, 80(4), 299–312. https://doi.org/10.1159/000511629

[1,2]Trickett, P. K., Gordis, E., Peckins, M. K., Susman, E. J., 2014. Stress reactivity in maltreated and comparison male and female young adolescents. Child Maltreatment, 19(1), 27–37. https://doi.org/10.1177/1077559513520466

[1]Tsai, N., Mukhopadhyay, S., Quas, J. A., 2021. Stress and working memory in children and adolescents: Insights from a multisystem approach. Journal of Experimental Child Psychology, 209, 105176. https://doi.org/10.1016/j.jecp.2021.105176

[1,2]Turan, B., Tackett, J. L., Lechtreck, M. T., Browning, W. R., 2015. Coordination of the cortisol and testosterone responses: A dual axis approach to understanding the response to social status threats. Psychoneuroendocrinology, 62, 59–68. https://doi.org/10.1016/j.psyneuen.2015.07.166

[1,2]Utge, S., Räikkönen, K., Kajantie, E., Lipsanen, J., Andersson, S., Strandberg, T., Reynolds, R. M., Eriksson, J. G., Lahti, J., 2018. Polygenic risk score of SERPINA6/SERPINA1 associates with diurnal and stress-induced HPA axis activity in children. Psychoneuroendocrinology, 93, 1–7. https://doi.org/10.1016/j.psyneuen.2018.04.009

[1,2]van Campen, J. S., Jansen, F. E., Pet, M. A., Otte, W. M., Hillegers, M. H. J., Joels, M., Braun, K. P. J., 2015. Relation between stress-precipitated seizures and the stress response in childhood epilepsy. Brain: A Journal of Neurology, 138(Pt 8), 2234–2248. https://doi.org/10.1093/brain/awv157

[1,2]Verdejo-Garcia, A., Moreno-Padilla, M., Carmen Garcia-Rios, M., Lopez-Torrecillas, F., Delgado-Rico, E., Schmidt-Rio-Valle, J., Fernandez-Serrano, M. J., 2015. Social stress increases cortisol and hampers attention in adolescents with excess weight. Plos One, 10(4), e0123565. https://doi.org/10.1371/journal.pone.0123565

[1]Wade, M., Sheridan, M. A., Zeanah, C. H., Fox, N. A., Nelson, C. A., McLaughlin, K. A., 2020. Environmental determinants of physiological reactivity to stress: The interacting effects of early life deprivation, caregiving quality, and stressful life events. Development and Psychopathology, 32(5), 1732–1742.

[1]Wadsworth, M. E., Ahlkvist, J. A., Jones, D. E., Pham, H., Rajagopalan, A., Genaro, B., 2022. Targeting the proximal mechanisms of stress adaptation in early adolescence to prevent mental health problems in youth in poverty. Journal of Clinical Child and Adolescent Psychology51(3), 344–359. https://doi.org/10.1080/15374416.2022.2073235

[1,2]Wadsworth, M. E., Bendezú, J. J., Loughlin-Presnal, J., Ahlkvist, J. A., Tilghman-Osborne, E., Bianco, H., Rindlaub, L., Hurwich-Reiss, E., 2018. Unlocking the black box: A multilevel analysis of preadolescent children's coping. Journal of Clinical Child and Adolescent Psychology 47(4), 527–541. https://doi.org/10.1080/15374416.2016.1141356

[1,2]Weyn, S., Van Leeuwen, K., Pluess, M., Goossens, L., Claes, S., Bosmans, G., Van Den Noortgate, W., Lutin, E., Brohl, A. S., Chubar, V., Geukens, F., Bijttebier, P., 2022. Individual differences in environmental sensitivity at physiological and phenotypic level: Two sides of the same coin? International Journal of Psychophysiology, 176, 36–53. https://doi.org/10.1016/j.ijpsycho.2022.02.010

[1]White, G. W., Rhind, S. G., Nathan, P. C., Caterini, J. E., Jones, H. N., Wells, G. D., 2022. Physiological stress reactivity in pediatric cancer survivors treated with chemotherapy. Pediatric Blood & Cancer, 69(11), e29881. https://doi.org/10.1002/pbc.29881

[1,2]Wijnant, K., Klosowska, J., Braet, C., Verbeken, S., De Henauw, S., Vanhaecke, L., Michels, N., 2021. Stress responsiveness and emotional eating depend on youngsters' chronic stress level and overweight. Nutrients, 13(10), 3654. https://doi.org/10.3390/nu13103654

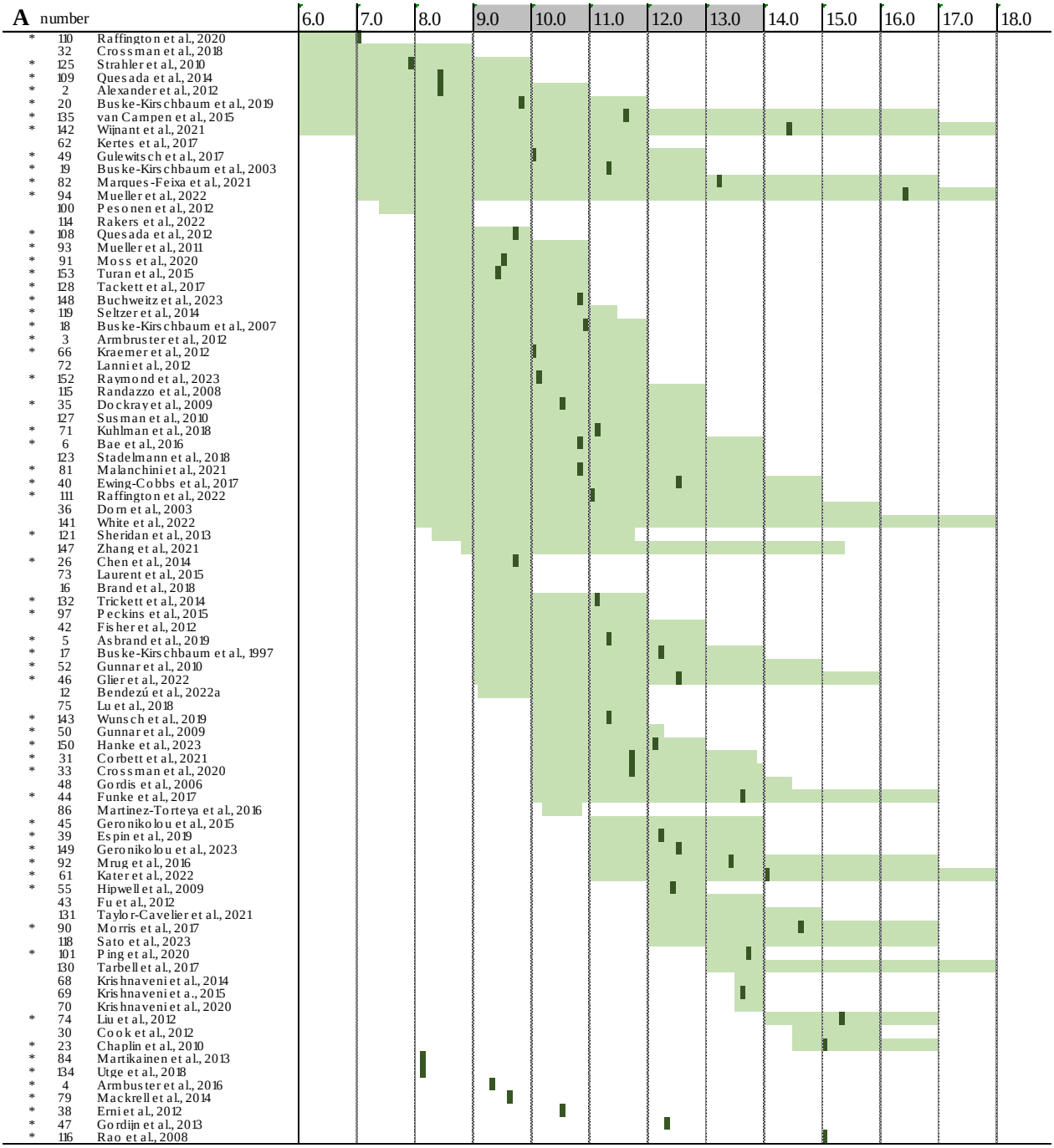
[1,2]Wunsch, K., Meier, M., Ueberholz, L., Strahler, J., Kasten, N., 2019. Acute psychosocial stress and working memory performance: The potential of physical activity to modulate cognitive functions in children. BMC Pediatrics, 19(1), 271. https://doi.org/10.1186/s12887-019-1637-x

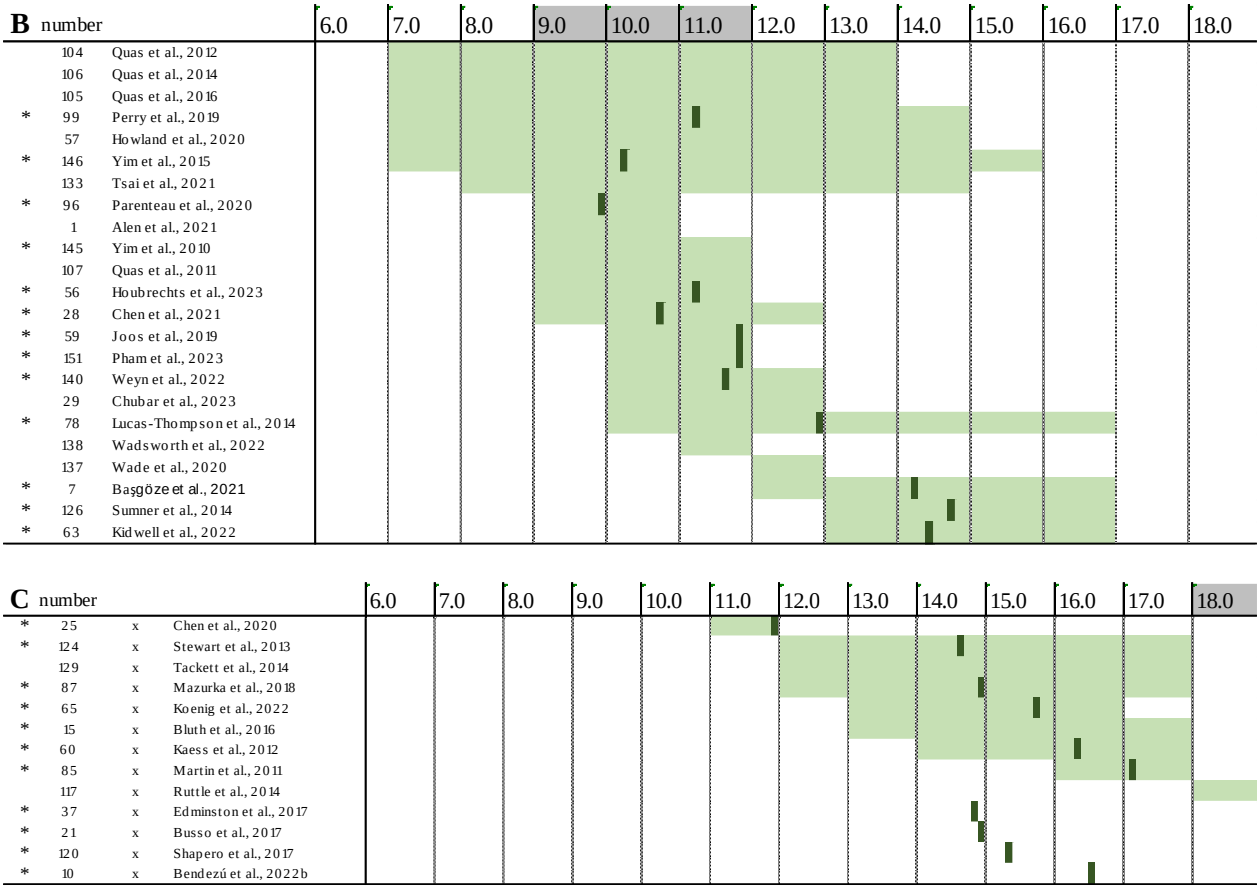
[1,2]Yim, I.S., Quas, J.A., Cahill, L., Hayakawa, C.M., 2010. Children's and adults' salivary cortisol responses to an identical psychosocial laboratory stressor. Psychoneuroendocrinology 35, 241–248. https://doi.org/10.1016/j.psyneuen.2009.06.014

[1,2]Yim, I.S., Quas, J.A., Rush, E.B., Granger, D.A., Skoluda, N., 2015. Experimental manipulation of the Trier Social Stress Test-Modified (TSST-M) to vary arousal across development. Psychoneuroendocrinology 57, 61–71. https://doi.org/10.1016/j.psyneuen.2015.03.021

[1]Zhang, D.-D., Fang, J., Zhang, L., Yuan, J.-Y., Wan, Y.-H., Su, P.-Y., Tao, F.-B., Sun, Y., 2021. Pubertal recalibration of cortisol reactivity following early life parent-child separation. Journal of Affective Disorders, 278, 320–326. https://doi.org/10.1016/j.jad.2020.09.030

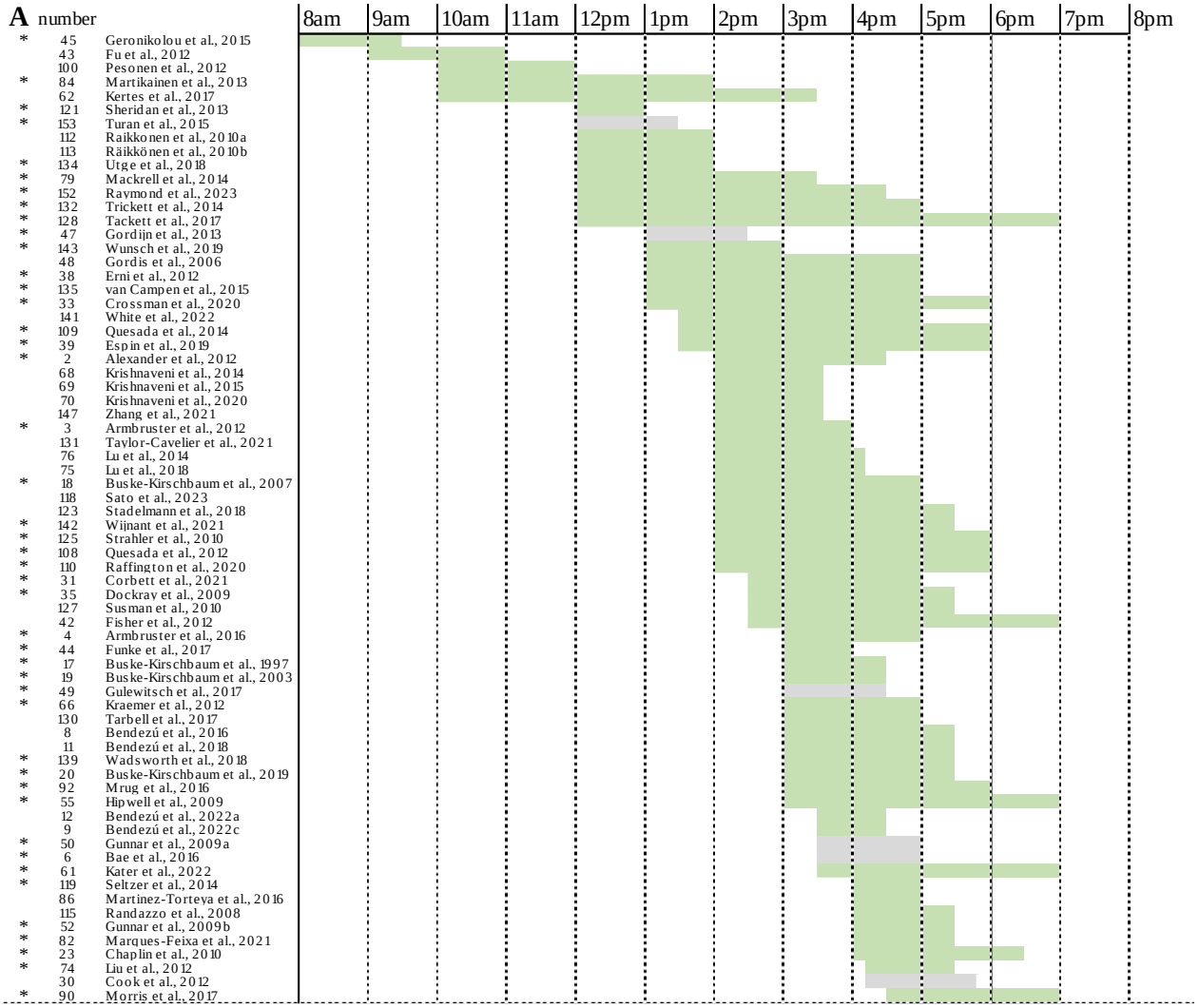
Figure S1. Age ranges applied in studies using (A) the TSST-C, (B) the TSST-M and (C) the original TSST procedure.

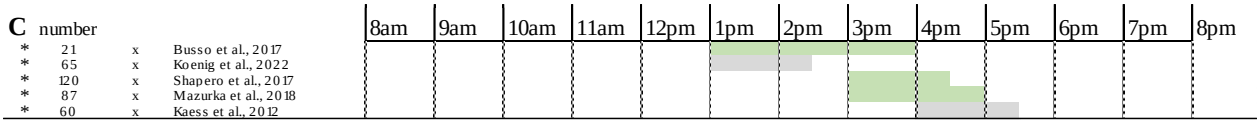




Note. Specified age ranges displayed in light green with sample age mean in dark green. Age range used in the validation study marked in dark grey. Studies that didn't report on an age range or a mean age for the relevant subsample (for TSST-C *k*=10, for TSST-M *k*=2) are not included. (*) marks studies that were included into the meta-analysis.

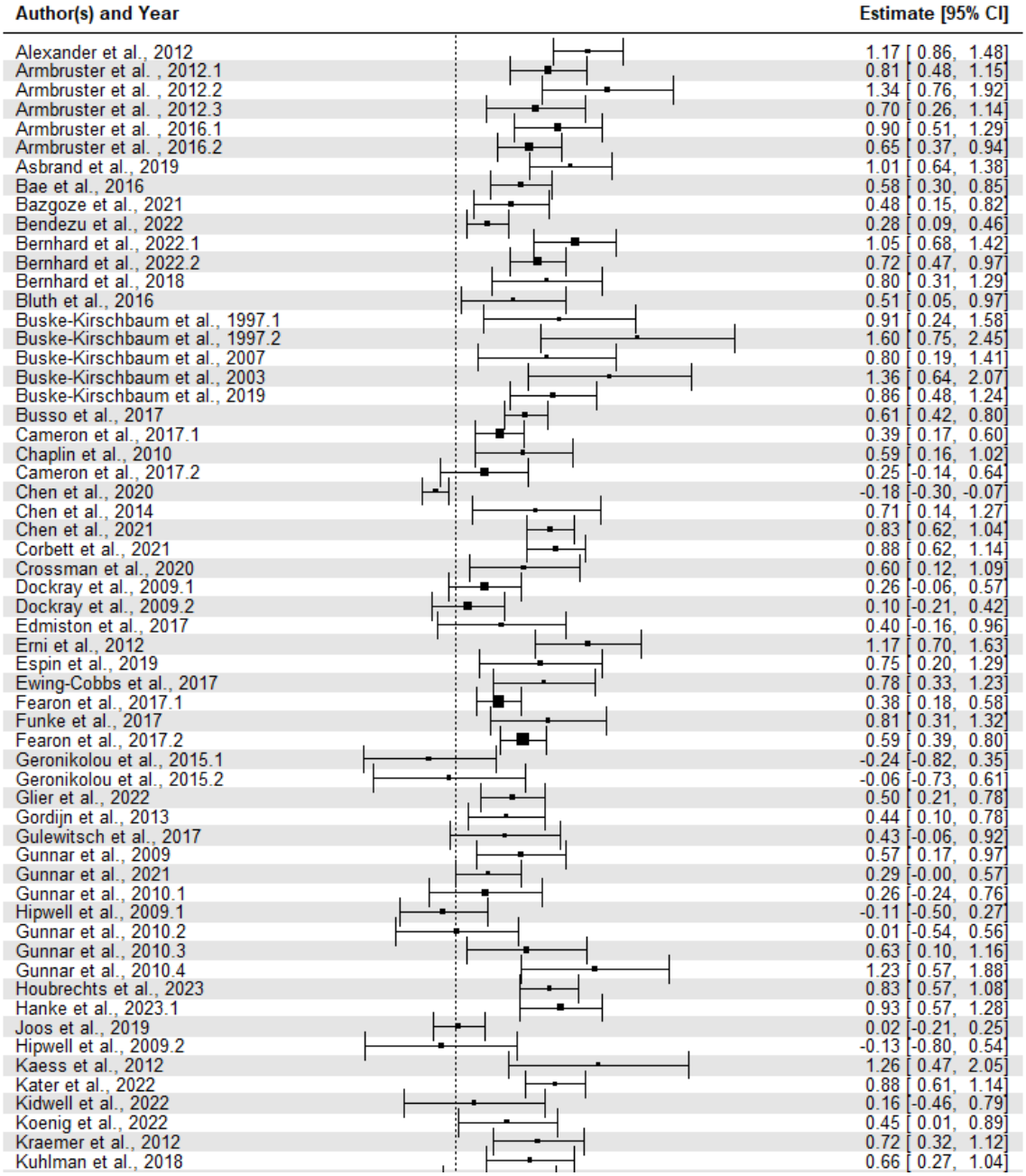
Figure S2. Time windows for assessments applied in studies using (A) the TSST-C, (B) the TSST-M, and (C) the original TSST protocol.





Note. Specified time windows displayed in light green; time windows specified only by starting time display in grey with one standard time window of 90 minutes. Studies that didn't report on a time window or a starting time (for TSST-C *k*=24, for TSST-M *k*=5, for TSST *k*=8) are not included. (*) marks studies that were included into the meta-analysis.

Figure S3. Forest plot of included studies and samples.



Continuation Figure S3. Forest plot of included studies and samples.

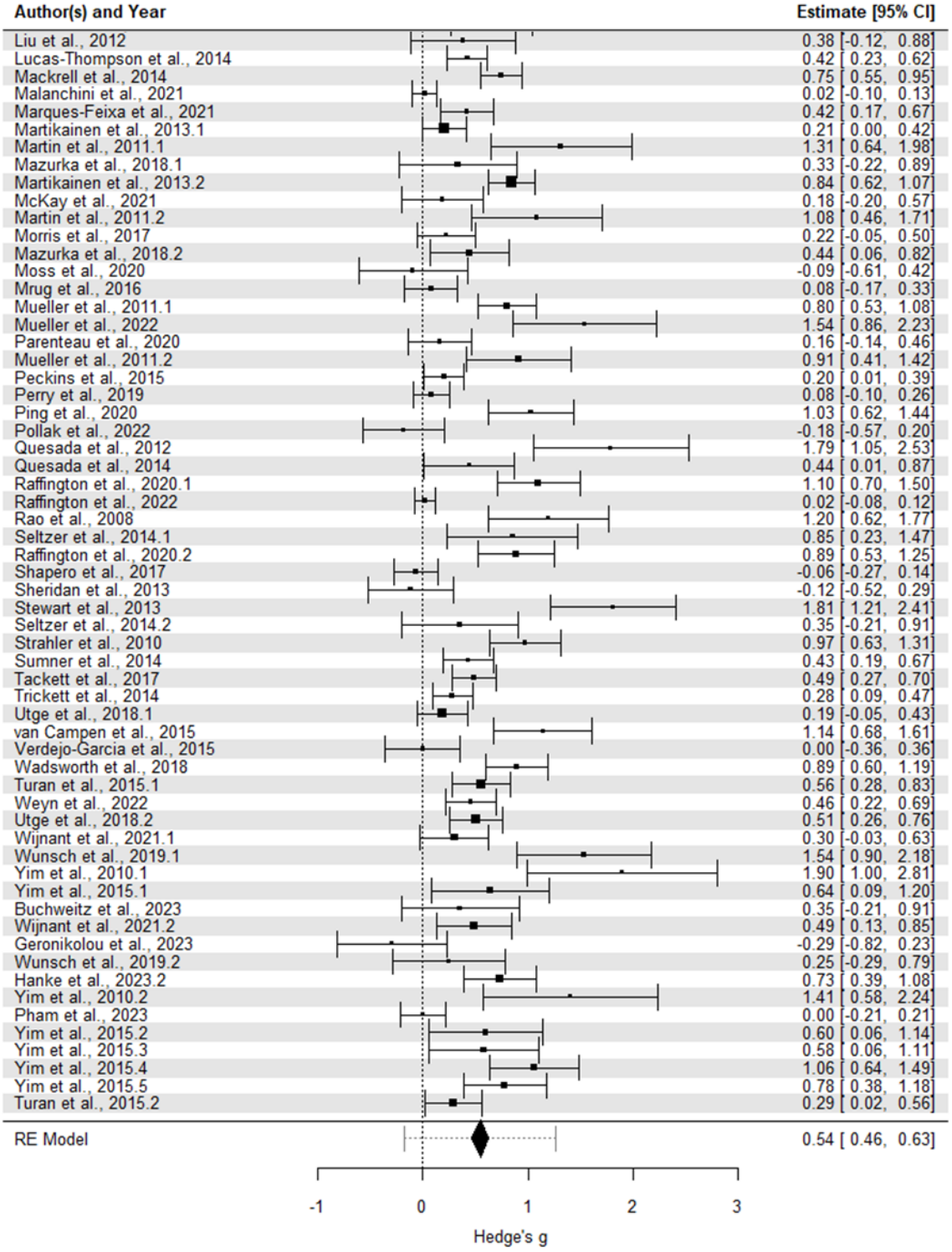
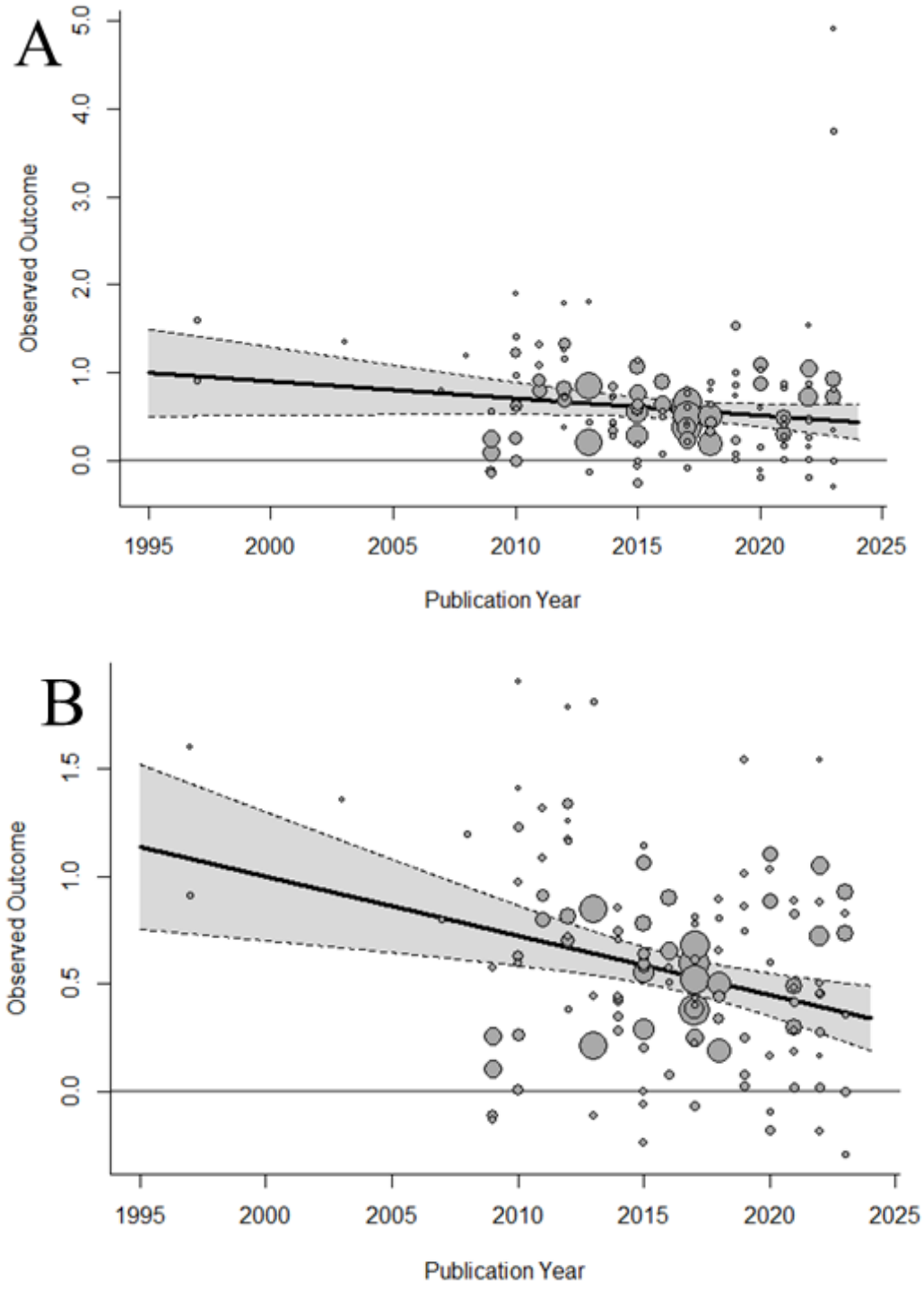
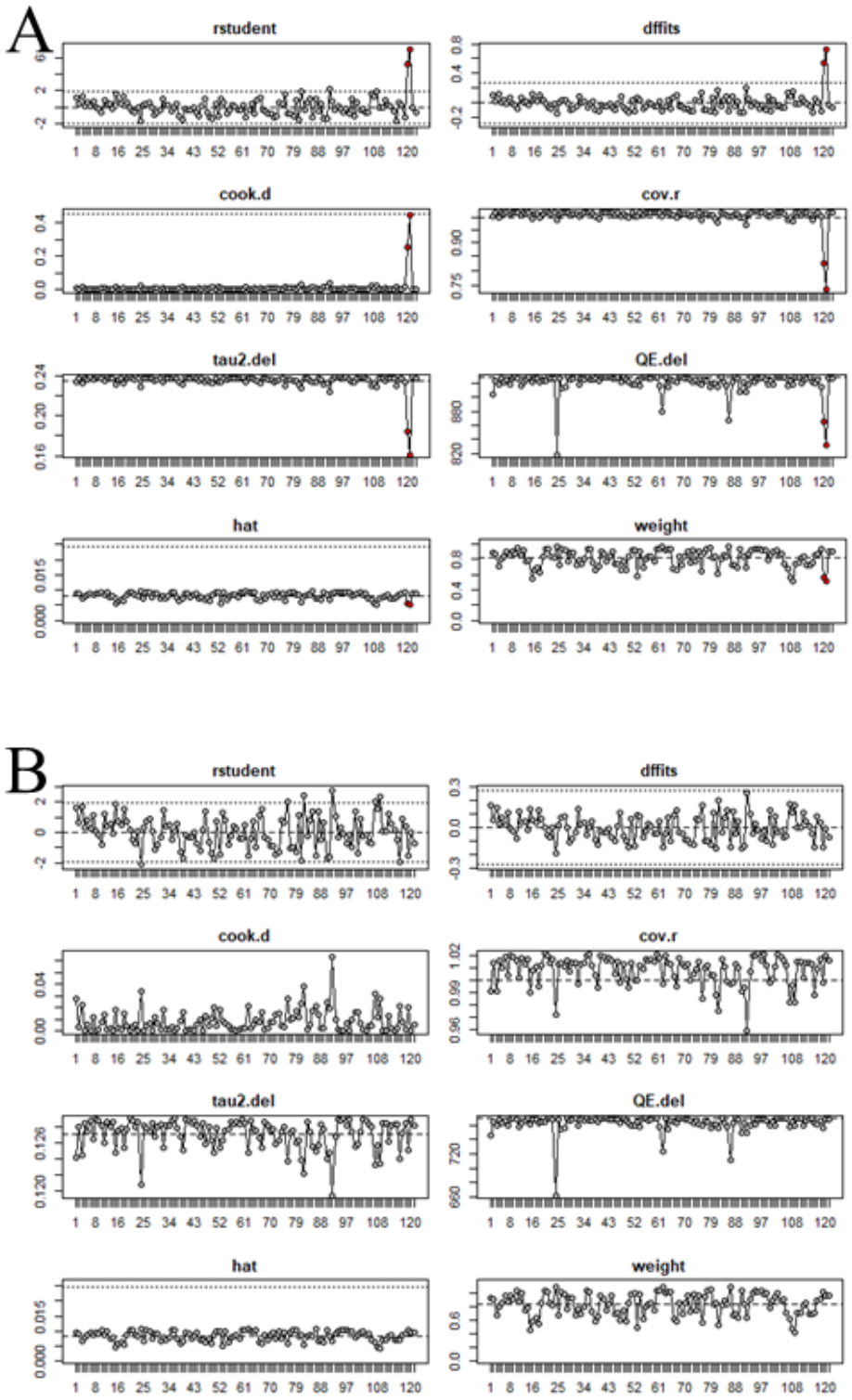


Figure S4. Bubble plot of the publication year bias regression



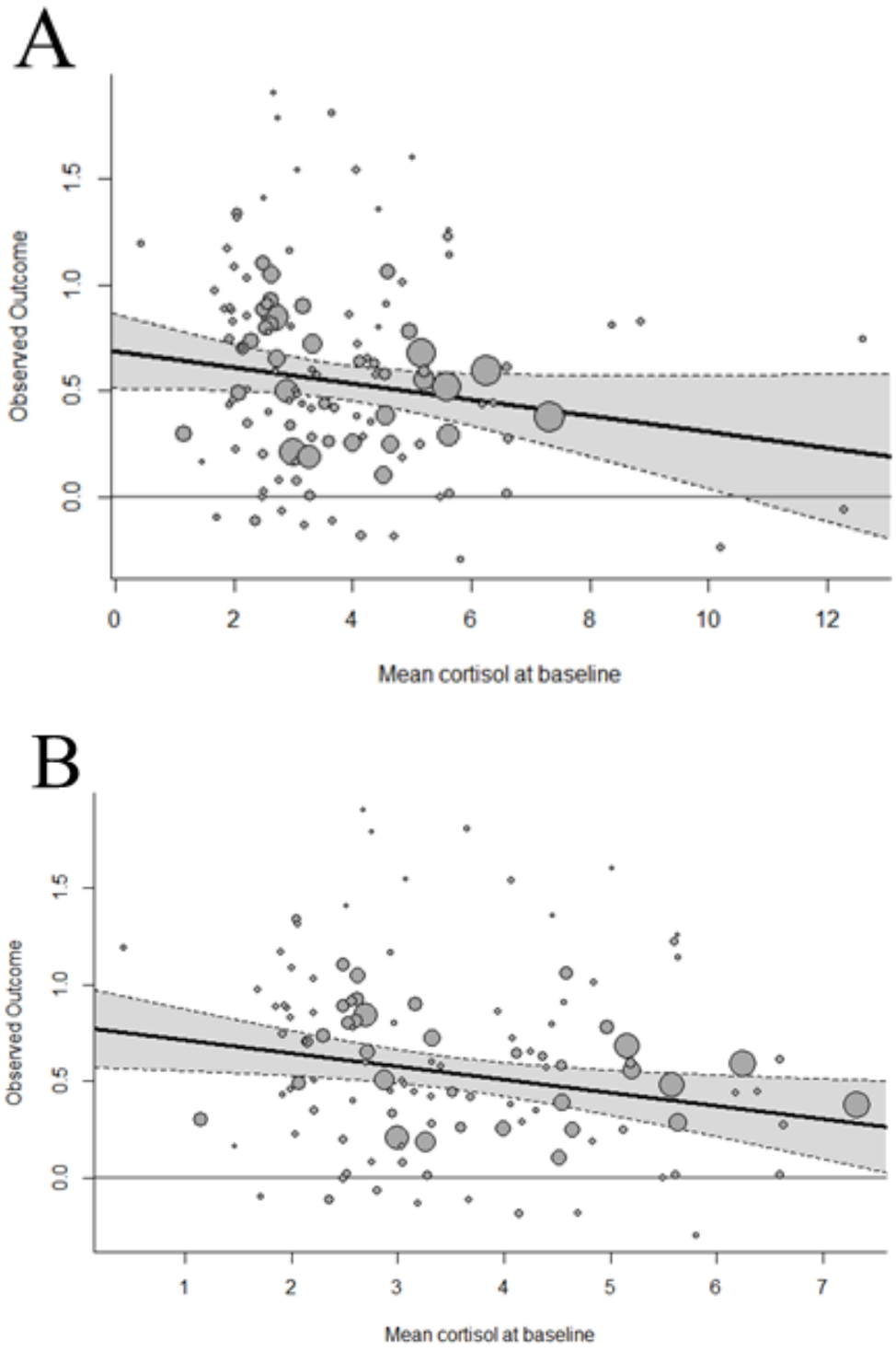
Note: Bubble plot 4a) without removing the influential effect sizes and 4b) after removing the influential effect sizes.

Figure S5: Influence plot.



Note: Influence plot 5a) before removing the influential effect sizes and 5b) after removing them.

Figure S6. Bubble plot for the moderation by baseline cortisol values.



Note: Bubble plot of the moderation by baseline cortisol 6a) before removing extreme values and 6b) after removing them.

Table S1. Exclusion criteria used in studies included in the qualitative analysis.

	TSST-C (Buske-Kirschbaum et al., 1997)	TSST-M (Yim et al., 2010)	TSST (Kirschbaum et al., 1993)
	N = 94	*N* = 23	*N* = 13
Exclusion Criteria			
Exclusion criteria defined			
Defined for participants	70 (74.5%)	16 (69.6%)	10 (76.9%)
Defined for parents	7 (7.5%)	0	0
Exclusion criteria in detail (participant-centric; percentage in relation to defined ones)			
Language skill in the language used	9 (12.9%)	2 (12.5%)	4 (40.0%)
Mental health issues (e.g., acute, or chronic mental health issues)	34 (48.6%)	8 (50.0%)	7 (70.0%)
Developmental disorders (e.g., Autism)	16 (22.9%)	7 (43.8%)	6 (60.0%)
Sensoric impairment (e.g., vision disorders)	3 (4.3%)	2 (12.5%)	0
Cognitive impairment	20 (28.6%)	5 (31.3%)	2 (20.0%)
Neurological disorders (e.g., epilepsy)	10 (14.3%)	1 (8.3%)	4 (40.0%)
Medical health (e.g., diabetes or acute illness)	39 (55.7%)	9 (56.3%)	5 (50.0%)
Research-specific	8 (11.4%)	1 (6.3%)	2 (20.0%)
Psychosocial issues (e.g., abuse)	2 (2.9%)	0	1 (10.0%)
HPA-affecting medication	45 (64.3%)	9 (56.3%)	3 (30.0%)
Specific medication	25 (35.7%)	4 (25.0%)	0
Steroid medication	28 (40.0%)	7 (43.8%)	2 (20.0%)
Other reasons	10 (14.3%)	4 (25.0%)	3 (30.0%)
Sex	0	0	1 (10.0%)
Pregnancy	1 (1.4%)	1 (6.3%)	3 (30.0%)
Birth complications	4 (5.7%)	0	0
Exclusion criteria in detail (parent-centric; percentage in relation to defined ones)			
Drugs during pregnancy	3 (42.9%)	0	0
Mental illness	2 (28.6%)	0	0
Chronic medical illness	1 (14.3%)	0	0
Language skills	2 (28.6%)	0	0

Table S2. TSST specifics.

	TSST-C (Buske-Kirschbaum et al., 1997)	TSST-M (Yim et al., 2010)	TSST (Kirschbaum et al., 1993)
	N = 94	*N* = 23	*N* = 13
	n (%) or *M(SD)* & (min - max)		
Used for age classes			
Children (6-12 years old)	32 (34.1%)	9 (39.1%)	1 (7.7%)
Mixed (6-18 years old)	33 (35.1%)	11 (47.8%)	0
Adolescents (12-18 years old)	13 (13.8%)	3 (13.0%)	9 (69.2%)
Not specified	16 (17.0%)	0	3 (23.1%)
Assessed in time window			
Morning (8am – 12pm)	3 (3.2%)	0	0
Morning – Lunch (8am – 2pm)	1 (1.1%)	0	0
Lunch time (12pm – 2pm)	4 (4.3%)	0	0
Morning-Afternoon (8am–5pm)	1 (1.1%)	0	0
Lunch–Afternoon (12pm–5pm)	12 (12.8%)	8 (34.8%)	2 (15.4%)
Afternoon (2pm – 5pm)	28 (29.8%)	6 (26.1%)	5 (38.4%)
Lunch–Evening (12pm – >5pm)	4 (4.3%)	1 (4.3%)	
Afternoon-Evening (>2pm)	21 (22.3%)	4 (17.4%)	
Unclear time window	20 (21.3%)	4 (17.4%)	6 (46.2%)
Phase duration (in minutes)			
Minimum acclimation time	*n* = 82 34.82 (±20.97) 0 – 120	*n* = 20 29.50 (±16.53) 5 – 70	*n* = 11 21.36 (±14.85) 10 – 60
Preparation phase	*n* = 82 5.17 (±1.67) 2 – 10	*n* = 22 4.77 (±2.33) 3 – 10	*n* = 12 5.42 (±2.35) 2 – 10
Speech task	*n* = 84 4.89 (±0.41) 3 – 6	*n* = 22 5.32 (±0.48) 5 – 6	*n* = 12 4.42 (±1.16) 2 – 5
Arithmetic task	*n* = 84 4.92 (±0.35) 3 – 5	*n* = 22 4.68 (±0.48) 4 – 5	*n* = 12 4.42 (±1.38) 1 – 5
Total TSST duration	*n* = 81 14.99 (±1.88) 9 – 20	*n* = 22 14.77 (±2.33) 13 – 20	*n* = 12 14.25 (±4.07) 6 – 20
Details about the judges			
Number of judges			
1	5 (5.3%)	1 (4.3%)	1 (7.7%)
2	55 (58.5%)	20 (87.0%)	10 (76.9%)
3	3 (3.2%)	0	0
4	1 (1.1%)	0	0
not specific	23 (24.5%)	2 (8.7%)	2 (15.4%)

Gender of judges			
All female	1 (1.1%)	2 (8.7%)	0
All male	0	0	0
Mixed	5 (5.3%)	8 (34.8%)	2 (15.4%)
Opposite sex to participant	0	0	0
Not specified	88 (93.6%)	13 (56.5%)	11 (84.6%)
Facial and behavioral feedback of the judges			
Neutral	30 (31.9%)	14 (60.9%)	5 (38.5%)
Positive / friendly	3 (3.2%)	0	1 (7.7%)
Negative / hostile	0	0	0
Progressively negative	0	0	0
Not specified	61 (64.9%)	9 (39.1%)	7 (53.8%)
Details about the experimenter			
Gender of the experimenter			
Same as participant	0	0	0
Different from participant	0	0	0
Mixed	0	0	0
Always female presenting	0	3 (13.0%)	0
Always male presenting	0	0	0
Not specified	94 (100%)	20 (87.0%)	13 (100%)
Interaction behavior of the experimenter			
Neutral	3 (3.2%)	2 (8.7%)	0
Positive / friendly	0	0	0
Negative / hostile	0	0	0
Not specified	91 (96.8%)	21 (91.3%)	13 (100%)
Is experimenter part of the jury?			
Yes	8 (8.5%)	2 (8.7%)	0
No	27 (28.7%)	8 (34.8%)	4 (30.8%)
Not specified	59 (62.8%)	13 (56.5%)	9 (69.2%)

Table S3. Saliva sampling methods used.

	TSST-C (Buske-Kirschbaum et al., 1997)	TSST-M (Yim et al., 2010)	TSST (Kirschbaum et al., 1993)
	N = 94	*N* = 23	*N* = 13
		N (%)	
Sampling method			
Salivette (Sarstedt. Germany)	50 (53.2%)	13 (56.5%)	4 (30.8%)
Passive drool	30 (31.9%)	9 (39.1%)	8 (61.5%)
SaliCaps (Tecan)	0	0	0
not specified	14 (14.9%)	1 (4.4%)	1 (7.7%)

Table S4. Reported responder rates and classification criteria used.

	TSST-C (Buske-Kirschbaum et al., 1997)	TSST-M (Yim et al., 2010)	TSST (Kirschbaum et al., 1993)
	N = 94	*N* = 23	*N* = 13
		N (%)	
Responder rate			
reported	6 (6.4%)	1 (4.3%)	1 (7.7%)
Criterium used to classify responders and non-responders			
1.5 nmol/L (Miller et al., 2013)	1 (16.7%)	0	0
2.5 nmol/L (Miller et al., 2013)	1 (16.7%)	0	0
≥15.5% baseline-to-peak increase (Miller et al., 2013)	0	1 (100%)	0
Positive AUCi	0	0	1 (100%)
≥10% baseline-to-peak increase (Ji et al., 2016)	1 (16.7%)	0	0
increase > 2x highest intra-assay coefficient of variance (Granger et al., 2012)	3 (50.0%)	0	0

Table S5. Reported responder rates.

			Control sample		Full sample	
Publication	TSST used	Criterium used	*n*	*%*	*n*	*%*
Asbrand et al., 2019	TSST-C	1.5 nmol/L	55	61.82	120	54.17
Espin et al., 2019	TSST-C	2.5 nmol/L	22	72.20	39	56.41
Fisher et al., 2012	TSST-C	Positive AUCi		NA	55	39.60
Trickett et al., 2014	TSST-C	≥10% increase	151	59.00	454	56.00
Mrug et al., 2016	TSST-C	≥10% increase		NA	84	36.00
Glier et al., 2022	TSST-C	≥10% increase		NA	80	80.00
Joos et al., 2019	TSST-M	≥15.5% increase		NA	101	11.00
Chen et al., 2020	TSST	≥2x highest intra-assay coefficient		NA	419	58.23
Gunnar et al., 2021	TSST-OL	1.5 nmol/L		NA	43	63.00
Bernhard et al., 2018	TSST (residual)	≥15.5% increase	28	82.00		NA

Table S6. Compiled comparison of publications included in quantitative analysis for Seddon et al. (2020) and Seel et al. (2024).

		Seel et al., 2024 (*N*= 92)	
		included	not included
Seddon et al., 2020 (*N*=57)	included	24	33 No cortisol data (*n*=9) Significant changes to TSST (*n*=7) No TSST (*n* =1) Grey literature (*n*=7) Only clinical sample (*n*=2) Raw data not supplied by authors (*n*=7)
	not included	68 Later publication (*n*=39) Reason unknown (*n*=29)	