Title
Modifying the Trier Social Stress Test to Induce Positive Affect

Journal
Affective Science, 2(4)

ISSN
2662-2041

Authors
Urban-Wojcik, Emily J
Charles, Susan T
Levine, Linda J

Publication Date
2021-12-01

DOI
10.1007/s42761-021-00074-6

Peer reviewed

Affective Science (2021) 2:427–437
https://doi.org/10.1007/s42761-021-00074-6

RESEARCH ARTICLE

Modifying the Trier Social Stress Test to Induce Positive Affect

Emily J. Urban-Wojcik[1] · Susan T. Charles[2] · Linda J. Levine[2]

Received: 5 January 2021 / Accepted: 17 August 2021 / Published online: 28 September 2021

Abstract
Studies comparing the effects of positive and negative affect on psychological outcomes are limited by differences in the situations that evoke these states and in the resulting levels of arousal. In the present research, we adapted the speech portion of the Trier Social Stress Test (TSST) to create conditions with similar situational features that induce either positive, negative, or neutral affective states ($N = 301$). Pre-post emotion ratings showed that negative affect increased in the negative condition but decreased in the positive and neutral conditions. Positive affect increased in the positive condition, remained unchanged in the neutral condition, and decreased in the negative condition. Participants' post-speech ratings of their positive and negative emotions differed significantly between the positive and negative conditions, which has not been accomplished in previous attempts to create a non-stressful positive TSST. Importantly, participants in the positive and negative conditions did not differ in self-reported levels of arousal and showed similar changes in mean arterial pressure across the speech period, although heart rate was relatively higher during the speech for participants in the negative compared to positive and neutral conditions. Findings demonstrate the effectiveness of a modified TSST for inducing positive affect with similar levels of emotional arousal to the traditional negative TSST.

Keywords Emotional valence · Positive affect induction · TSST

Introduction

Subjective emotional experiences influence a number of highly important psychological processes, such as attention (Compton, 2003), memory (Kensinger, 2009; Levine & Pizarro, 2004), and decision-making (Lerner et al., 2015). However, in most early experimental work examining these associations, researchers have either combined positive and negative stimuli into one category to compare emotionally-arousing to neutral stimuli or only compared negative to neutral stimuli. Thus, early studies were not designed to compare the effects of valence on psychological processes (Levine & Pizarro, 2004). It is now well-accepted that positive and negative affect are theoretically and methodologically separable dimensions (Russell & Carroll, 1999; Watson et al., 1988), with dissociable regions of activation in the brain (Dolcos et al., 2004; Kensinger & Schacter, 2006). Researchers acknowledge that positive and negative affect might also differentially impact psychological processes of interest (Kensinger, 2009), and more recent research directly compares the effects of positive to negative affect on a number of outcomes.

Despite this shift, current research contrasting the effects of valence is limited by the inability to control or equate the content of positive and negative emotional experiences. Situations and stimuli that elicit positive emotions are often inherently different from those that create negative emotions. As a result, the majority of research on the impact of emotional experiences on psychological and physiological outcomes focuses on negative emotional experiences. However, a growing number of studies are beginning to shed light on the powerful role that positive affect plays in shaping cognitions, behaviors, and health (e.g., Howell et al., 2007; Lyubomirsky et al., 2005; Sin et al., 2015; Talarico et al., 2009), and experimental manipulations are needed to directly test the influence of valence on these outcomes. For this reason, we sought to design an ecologically valid

Handling Editor: Michael Kraus

✉ Emily J. Urban-Wojcik
ejurban@wisc.edu

[1] Center for Healthy Minds, University of Wisconsin-Madison, Madison, WI, USA

[2] Department of Psychological Science, University of California, Irvine, Irvine, CA, USA

experimental affective state induction procedure that would allow for a controlled study of how emotional valence influences a number of important constructs while also eliminating the possible confounding influence of arousal. The speech portion of the TSST, a commonly used laboratory-based psychosocial stressor requiring participants to give an impromptu speech in front of stern evaluators, reliably elicits high-arousal negative affect under controlled conditions (Dickerson & Kemeny, 2004; Dickerson et al., 2008; Kirschbaum et al., 1993). A high-arousal, positive version of the TSST would provide researchers with an ecologically valid setting to contrast the influence of emotional valence on a number of outcomes while ensuring a similar setting between conditions.

Three groups of researchers have created more positive versions of the TSST. In each case, participants gave speeches on topics related to a job interview or their career aspirations in front of trained evaluators, similar to the classic TSST. In the positive conditions, however, evaluators responded positively to the participants non-verbally (e.g., smiling, nodding, and leaning forward) and in some cases verbally. The researchers differed in their goals for a positive TSST, however. Whereas some researchers aimed to create a positive stress condition (Akinola & Mendes, 2008; Crum et al., 2017; Kassam et al., 2009; Taylor et al., 2010; Waters et al., 2014), others attempted to create a non-stressful positive TSST (Herten et al., 2016; Wiemers & Wolf, 2015; Wiemers et al., 2013a, 2013b, 2013c). When physiological responses were assessed, there was evidence of physiological stress response in both the negative and positive conditions, as indexed by decreases in pre-ejection period (Kassam et al., 2009), greater ventricle contractility (Waters et al., 2014), increases in cortisol (Crum et al., 2017; Taylor et al., 2010), or increases in salivary alpha-amylase (Herten et al., 2016; Wiemers et al., 2013a, 2013b, 2013c) between baseline and during/after the speech. Although positive affect did increase in response to some iterations of the positive, or “friendly” versions of the TSST (Crum et al., 2017; Herten et al., 2016; Waters et al., 2014; Wiemers & Wolf, 2015), these increases did not always result in significant differences in post-speech positive affect between conditions.

These studies provide strong models for a positive TSST manipulation and have greatly contributed to our knowledge of how social situations, experienced in threatening, challenging, and supportive contexts, influence physiological reactivity and performance. They have provided insight into the complexities of affective responses to high-arousal social contexts, demonstrating that social stress can be experienced positively as challenge (e.g., Crum et al., 2017), that negative stressful experiences can nonetheless facilitate performance (e.g., Wiemers et al., 2013a, 2013b, 2013c), that physiological reactivity to a stressful experience is “contagious” between close dyads (Waters et al., 2014), and that responses to stressful social situations can vary as a function of individual differences (Akinola & Mendes, 2008).

In these previous iterations of a positive TSST, however, the task was still associated with a stress response, as measured by increases in negative affect or physiological reactivity. The overall aim of the present research, therefore, was to design a non-stressful, yet similarly arousing positive control manipulation that mirrored the classic negative TSST. We had two specific goals. Given that people typically experience negative events as more impactful than positive events (Baumeister et al., 2001), our first goal was to address the theoretical question of whether a non-stressful positive experience could be perceived to be as arousing as a stressful negative experience. To address this question, we diverged from previous positive TSST manipulations by creating a shorter TSST manipulation (previous manipulations lasted between 13 and 15 min) and making the speech topic in the positive condition more enjoyable. As with previous positive TSST iterations, we trained evaluators to be a supportive audience. Our second goal in designing this positive version of the TSST was to provide experimentalists with a method to compare the effects of positive and negative valence on a multitude of psychological outcomes while ensuring similarly high levels of arousal across conditions. Such a manipulation could be useful when examining the effects of valence on cognitive processes such as attention, decision making, and memory, as well as investigating how carefully matched positive compared to negative social interactions influence mental and physical health.

For the sake of developing a robust, positive emotional lab experience, we prioritized measures of self-reported positive and negative emotion in response to the manipulations. We reasoned that a positive TSST should first and foremost be subjectively experienced as positive, and in-depth assessment of participants’ subjective emotional state as well as their appraisals of the task was the most face valid way to develop a successful manipulation. Self-report is only one method for assessing emotional experiences, however, and it is useful to interpret self-report data in concert with physiological and behavioral measures as well. For this reason, we included cardiovascular measures of blood pressure (BP) and heart rate (HR) as a first glimpse into the physiological impact of our manipulation. Although these metrics do not provide a comprehensive assessment of autonomic system activation, they are useful in assessing general physiological arousal in emotion research (Cacioppo et al., 2000; Mauss & Robinson, 2009). We have made all of our evaluator instructional materials and detailed descriptions of each condition available to facilitate their implementation in future research. Study materials, computed variables, and analytic syntax are available at https://osf.io/hkvqs/?view_only=6336c6421cd8471aa855909efb1fa7d0.

Method

Participants

We used the pwr package in RStudio (pwr 1.2–2; Champely, 2018) to determine our sample size. In four previous TSST studies that included a positive condition and reported changes in negative affect (NA) from baseline to post-TSST, the weighted *r* effect size for condition differences in post-TSST NA was 0.43 (Herten et al., 2016; Wiemers et al., 2013a, 2013b, 2013c). These studies either did not measure change in positive affect (PA) or provide sufficient information to calculate the effect size for the PA change by condition interaction. We predicted that condition differences in PA after the TSST would be slightly smaller (i.e., a medium effect size around *r* of 0.30). For three conditions (negative TSST, neutral TSST, and positive TSST), 34 participants were required in each condition to detect a medium effect size at a significance level of $p = 0.05$ with power of 0.80.

Once the desired sample size was collected (35 in each condition), we collected data from a second, larger cohort of individuals to replicate the results and to assess physiological responses to the task by collecting repeated BP measurements. Thus, our total sample can be divided into two cohorts. Because speech task procedures were the same between cohorts (with the exception of the BP measurements in cohort 2), and results of the affective state induction were similar[1] across both cohorts, we concatenated data from both cohorts for the sake of brevity and power (see Supplementary Materials A for affective state induction results by cohort).

Cohort 1

The first cohort consisted of 105 eligible undergraduate students recruited from the University's human subject lab pool to participate for course credit. Participants were eligible for the study if they were at least 18 years of age, had a strong command of the English language, and were not currently taking mood-altering medications, such as anti-depressants. Participants ranged from 18 to 33 years old ($M = 20.76$, $SD = 2.48$) and 77% identified as female. Participants were 49.5% Asian/Pacific Islander, 29.5% Hispanic, 9.5% Caucasian, 2.9% African American, 6.7% Biracial/Multiracial, and 1.9% other ethnicity. More than two-thirds (70.5%) of the sample was born in the United States.

Cohort 2

Another 199 participants were recruited through the University's human subject pool. Two individuals asked that their data be withheld and one lacked baseline emotion data, resulting in a final sample of 196 individuals. Due to a technical error, demographic information was not collected during the session but was gathered from a pre-screen[2] filled out by each participant. More than three quarters (82.1%) of the sample identified as being female. Participants were 52.6% Asian/Pacific Islander, 25% Hispanic, 10.7% Caucasian, 1.5% African American, 7.1% Biracial/Multiracial, 2% other ethnicities, and 1% declined to answer. The majority (80.6%) were native English speakers. At the time of the pre-screen questionnaire, a total of 92.3% of participants indicated being 21 and younger and 97.4% indicated being 25 and younger. The median age range for participants was 18–21.

Procedure

After informed consent and the eligibility screening questionnaire, participants completed a computer-administered questionnaire assessing baseline emotions. TSST condition (positive, neutral, or negative) was randomly assigned prior to the session.[3] Two new researchers (the "evaluators") then administered the TSST procedure. The group of evaluators was similar in age, ethnicity, and gender to the pool of participants, although evaluators were not specifically assigned to work with certain participants based on these characteristics.

In each TSST condition (negative, positive, or neutral), the evaluators entered the room, gave the participants

[1] Only three minor differences can be noted regarding the affective state induction findings between cohorts 1 and 2. The neutral condition in cohort 2 had significantly higher baseline positive affect than the positive condition (but not the negative condition). The positive condition in cohort 2 saw a pre-post decrease in negative affect, but negative affect did not change for the positive condition of cohort 1. Finally, on the single item rating of negative valence, the neutral and positive conditions did not differ significantly for cohort 2, whereas they did for cohort 1. The remaining results (e.g., positive emotion increasing pre- to post- speech for the positive condition; positive emotion decreasing and negative emotion increasing for the negative condition; no positive or negative emotion changes for the neutral condition) followed a similar pattern for both cohorts.

[2] Participants completed the pre-screen anywhere between 3.5 years before the commencement of the current study to immediately before signing up to participate. Age was measured by asking participants to indicate their age within a set of age ranges (17 and under, 18–21, 22–25, 26–30, and 31–40). Because of these age ranges and the variable time lapse between when the pre-screen was completed and when the current study was completed, exact age for cohort 2 participants could not be computed.

[3] Given resource constraints during data collection for cohort 2, the final half of data collection ran a greater proportion of positive and negative trials than neutral trials. Because of this, sample sizes in the neutral condition are smaller than the positive and negative conditions.

instructions for the speech task (described below), a pen and a pad of paper, and answered any questions. After leaving the participant alone for two minutes to prepare, the evaluators then re-entered the room and participants began their speech. The speech task lasted five minutes and was video recorded.

Afterward, the evaluators instructed the participant to complete a set of questionnaires on the computer and left the room. These final questionnaires assessed the emotions they experienced during the TSST, their appraisals of the task, whether they had completed a similar task in the past, and a demographics questionnaire (for cohort 1 only). The original researcher then returned to debrief participants. All study procedures were approved by the university's institutional review board. More details about each condition as well as the evaluator scripts and instructions are available in Supplementary Materials B.

Negative TSST

The negative TSST condition was similar to the classic TSST speech task. Before the speech task, participants were instructed by the TSST evaluators to imagine they were interviewing for a demanding new job related to their career aspirations. Participants were told that the purpose of the exercise was to gather information for school officials on how prepared the university's students were for the job market and that the evaluators had been trained to analyze the participant's performance. In addition, they were told that a video recording of their speech would be evaluated by a separate panel for style, eloquence, and overall quality. They were allowed to make notes during the preparation period but were told that they could not use them during the speech task.

During the speech, the trained evaluators wore white lab coats, maintained a neutral expression, and took notes on a clipboard. They refrained from nodding, smiling, or giving any words of encouragement. Their posture was stiff and reserved. Approximately every 45 s, evaluators took turns interjecting with scripted comments such as "You are spending too much time on this aspect." After five minutes, the evaluators abruptly interrupted the speech (if the participant was still talking) and stated that the speech portion of the session was completed before leaving the room.

Positive TSST

The positive TSST condition was as similar to the negative condition as possible but designed to increase positive emotion and decrease negative emotion. The two TSST evaluators introduced themselves warmly upon meeting the participant to give speech instructions. They instructed the participant to imagine that they were being interviewed by two new university students interested in starting a new hobby and to describe one of their favorite hobbies or activities and why they enjoy it so much. They were told that the purpose of this exercise was to gather information for incoming students on what current students are passionate about doing. Participants were told that they were not being evaluated and that the video recording was meant to help the researchers confirm notes after the session. Participants then had two minutes to prepare for the speech using a pen and notepad, which they could use during the speech.

During the speech, the evaluators mirrored the affect of the participant but also displayed compassionate and understanding facial expressions. Evaluators were not wearing lab coats, maintained comfortable eye contact, and gave verbal and non-verbal signals of engagement (e.g., nodding, laughing at appropriate moments, smiling genuinely, and saying "uh-huh"). Their posture was attentive and relaxed. Evaluators were instructed to avoid awkward pauses if possible and took turns making scripted comments or questions (approximately every 45 s). Comments expressed interest (e.g., "Wow, that's really cool.") and questions were meant to help avoid awkward pauses (e.g., "How did you get started with this hobby?"). After five minutes, the evaluators indicated that the speech period was finished and that they thought the participant gave a lot of good information regarding the participant's hobby.

Neutral TSST

The neutral TSST was designed to provide a control condition for the positive and negative conditions. TSST evaluators told the participant to imagine they were giving a speech to high school students interested in attending the university about the details of their daily routine during a typical weekday from the time they awoke. Participants were told the purpose of this exercise was to gather information for local high school students on the daily routines of university students. Similar to the positive condition, participants were told that they were not being evaluated and that the purpose of the video recording was to help the researchers confirm notes after the session. Participants had two minutes to prepare for the speech using a pen and notepad, which they were allowed to use during the speech.

Evaluators of the neutral condition did not wear lab coats but maintained a professional demeanor. They were kind but refrained from laughing or smiling genuinely. Evaluators were allowed to nod, smile socially (i.e., without crinkling the corners of their eyes), and take notes to show engagement. They did not make any comments or question unless an awkward pause occurred, in which case evaluators took turns probing for more details (e.g., "What do you do next?", "Can you tell me about this in a little bit more detail?").

After five minutes, the evaluators ended the speech and thanked the participant for their involvement in the study.

Measures

Baseline and Post-TSST Positive and Negative Affect

Participants rated how accurately 26 emotion items (12 positive, 14 negative) described how they were currently feeling at the beginning of the study and again after the speech task on a scale of 1 "Not at all Accurate" to 5 "Extremely Accurate." Items were chosen from the Profile of Mood States (POMS; McNair et al., 1971) and the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) and also included researcher-generated items (see Supplementary Materials C). The scale was designed to include a similar number of low and high arousal positive and negative emotions. Positive items ($\alpha = 0.90$) and negative items ($\alpha = 0.85$) were averaged separately for each participant to allow for tests of change in emotional valence by condition.

Post-TSST Overall Valence and Arousal

After the speech task, participants rated how positive and how negative the task made them feel overall on two separate scales from 1 "Not at all/Neutral" to 7 "Very Positive" and 1 "Not at all/Neutral" to 7 "Very Negative." If participants indicated feeling at least some positive emotion on the previous measure (i.e., a 2 or higher), they were also asked to rate how activating or arousing the positive emotions were that they felt during the speech task overall, using a scale of 1 "Not at all/I felt calm" to 7 "Very much/I felt excited." The same procedure was used for negative emotions, using a scale of 1 "Not at all/I felt bored" to 7 "Very much/I felt tense."

Appraisals

A series of 20 additional questions assessed appraisals of the TSST experience (see Supplementary Materials D). These questions were adapted from the Performance Attribution Questionnaire (Dickerson et al., 2009), which has been used in previous TSST research. All questions were assessed on a 1 to 7 scale except for the question regarding friendliness of the TSST researchers, which was assessed on a scale of 0 to 100 but transformed to match the same 1 to 7 scale. As a manipulation check, two additional open-ended questions asked participants to state what they thought the goal of the task was according to the researcher and what their own goal was during the task. We included these questions to gain a more qualitative understanding of differences between the conditions from the participant's point of view. A principal components analysis using an oblique rotation was computed to group similar questions. This analysis resulted in four underlying components, described below. Items within components were combined to create a composite score for each component (after reverse-coding negatively loading items). Composite scores were then statistically compared between conditions using four one-way ANOVAs.

Four components (Eigenvalues of 10.47, 1.64, 1.27, and 1.21) accounted for 72.91% of the variance in appraisals (52.33%, 8.18%, 6.36%, and 6.04%, respectively). The first component, labeled positive self-performance, was composed of nine items. Higher scores indicated a more favorable self-performance rating, with appraisals of meeting the researcher's and the individual's own goal for the task, rating their performance highly, predicting that others would rate their performance highly, having more confidence and control, reporting less task difficulty, and giving up or wanting to give up to a lesser extent. The second component, labeled challenge, consisted of three items where higher scores indicated finding the speech task more stressful and challenging and exerting less effort. The third component, labeled perception of evaluation, captured how much participants perceived that their speech was being evaluated and was measured by two items (speech was evaluated during the task, speech could be evaluated later). Higher scores indicated a stronger belief that their speech was or could be evaluated. The final component, labeled negative social evaluations, was composed of six items; higher scores on this composite indicated feeling more disliked, rejected, threatened, feeling less liked and accepted, and evaluating TSST administrators as less friendly.

Physiological Arousal

Participants in cohort 2 additionally were affixed with a BP cuff. Six measurements using a GE CARESCAPE V100-1 BP machine were taken every 60 s during the baseline and recovery periods with the first measurement taken at 00 s and the last measurement taken at 5 min. In addition, five measurements were taken every 60 s during the speech-task portion, where the first measurement was taken at 30 s and the last was taken at 4 min and 30 s. Each measurement produced a reading of systolic blood pressure (SBP), diastolic blood pressure (DBP), mean arterial pressure (MAP), and heart rate (HR). For each segment of the study (baseline, during the speech-task, and during recovery), the average values for level of SDP, DBP, MAP, and HR were calculated for each participant. Given that MAP is computed as a function of SDP and DBP, only results describing MAP and HR are reported for the sake of parsimony. Results specific to SBP and DBP are available in Supplementary Materials E.

Table 1 Affect means, standard errors, and 95% CIs by valence and condition

	Baseline positive affect		Post-speech positive affect	
Condition	*M* (*SE*)	95% CI	*M* (*SE*)	95% CI
Positive	2.55 (0.07)	[2.41, 2.68]	2.94 (0.08)	[2.79, 3.09]
Neutral	2.74 (0.10)	[2.55, 2.92]	2.69 (0.10)	[2.49, 2.90]
Negative	2.64 (0.07)	[2.50, 2.78]	2.24 (0.08)	[2.09, 2.39]
	Baseline negative affect		Post-speech negative affect	
Condition	*M* (*SE*)	95% CI	*M* (*SE*)	95% CI
Positive	2.03 (0.06)	[1.91, 2.15]	1.73 (0.07)	[1.60, 1.86]
Neutral	1.98 (0.08)	[1.82, 2.14]	1.80 (0.09)	[1.63, 1.98]
Negative	2.07 (0.06)	[1.95, 2.19]	2.43 (0.07)	[2.30, 2.55]

Note: sample sizes of each condition: positive ($n = 117$), neutral ($n = 64$), and negative ($n = 120$)

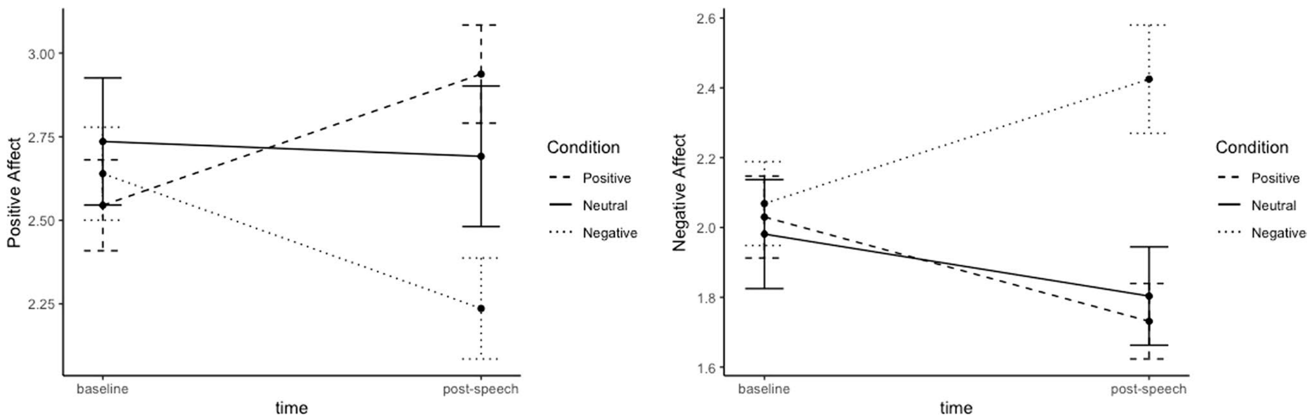

Fig. 1 Change in positive and negative affect by condition. Error bars represent 95% confidence intervals

Analyses

We conducted repeated-measures ANOVAs to test condition differences (positive vs. neutral vs. negative) in change in positive and negative affect and BP as well as one-way ANOVAs for items that were only measured once (e.g., single-item rating of subjective arousal). For all significant interactions, condition means were compared using 95% confidence intervals (repeated measures ANOVAs) and planned contrast tests (one-way ANOVAs) to determine the nature of the significant effect. In the case of a violation to the assumption of homogeneous variances, contrast statistics used corrected degrees of freedom (Games-Howell for one-way and Greenhouse–Geisser for repeated measures).

Results

Affective State Induction Efficacy

We conducted two repeated measures ANOVAs to assess change in positive and negative affect between baseline and just after the speech task. In each model, time (baseline vs. post-speech) was the within subject factor and condition (positive, neutral, negative) was the between-subjects factor. Table 1 displays means, standard errors, and 95% confidence intervals by valence and condition, and Fig. 1 exhibits a graphical representation.

For positive affect, a main effect of condition, $F(2, 298) = 5.62$, $p = 0.004$, $\eta_p^2 = 0.04$, and a significant condition by time interaction were found, $F(2, 298) = 54.63$, $p < 0.001$, $\eta_p^2 = 0.27$. Baseline positive affect did not differ significantly for participants in the positive and negative conditions, but participants in the neutral condition started with significantly higher positive affect than the positive condition (see Table 1 for confidence intervals). After the speech task, and adjusting for baseline positive affect, the positive condition reported significantly more positive affect than the neutral condition, which in turn reported significantly more positive affect than the negative condition. Between baseline and post-speech, paired *t*-tests demonstrated that average positive affect increased for the positive condition, $t(116) = 6.80$, $p < 0.001$, $d = 0.63$, decreased for the negative condition $t(119) = -7.73$, $p < 0.001$, $d = 0.71$, and did not change for the neutral condition, $t(63) = -0.66$, $p = 0.52$, $d = 0.08$.

For change in negative affect over time, a significant main effect of condition, $F(2, 298) = 12.26$, $p = 0.001$, $\eta_p^2 = 0.08$, and an interaction between condition and time were found, $F(2, 298) = 41.66$, $p < 0.001$, $\eta_p^2 = 0.22$. Baseline negative affect was similar across conditions. As Table 1 shows, participants in the negative condition reported significantly more post-task negative affect than the neutral and positive conditions. For participants in the negative condition, negative affect increased significantly from baseline levels, $t(119) = 5.58$, $p < 0.001$, $d = 0.51$. The positive and neutral conditions did not differ significantly from one another post-speech, but both significantly decreased in negative affect between baseline and post-speech, positive: $t(116) = -6.70$, $p < 0.001$, $d = 0.62$; neutral: $t(63) = -3.06$, $p = 0.003$, $d = 0.38$.

Single-Item Valence and Arousal

Two one-way ANOVAs were used to examine the single-item assessments of how positive or negative participants thought the speech task was overall. A significant omnibus test, $F(2, 298) = 84.05$, $p < 0.001$, $\eta^2 = 0.36$, and planned contrasts revealed that participants in the positive condition reported the speech task to be significantly more positive overall than the neutral condition, $t(108.10) = 3.65$, $p < 0.001$, and negative condition, $t(234.98) = 13.52$, $p < 0.001$, and that the neutral condition also found the task more positive than the negative condition, $t(110.28) = 6.20$, $p < 0.001$. For ratings of overall negativity by condition, $F(2, 298) = 117.13$, $p < 0.001$, $\eta^2 = 0.44$, participants in the negative condition found the speech significantly more negative than the neutral condition, $t(166.14) = 10.09$, $p < 0.001$, and positive condition, $t(187.04) = 14.24$, $p < 0.001$. Participants in the neutral condition rated the speech task more negatively than the positive condition, $t(104.65) = 1.90$, $p = 0.06$, although this effect did not reach statistical significance.

A total of 251 participants rated the speech task at least a little positive using the single-item measure (116 in the positive condition, 57 in the neutral condition, and 78 in the negative condition), and 199 participants found the task at least a little negative (53 in the positive condition, 34 in the neutral condition, and 112 in the negative condition). Those who reported finding the task at least a little positive or at least a little negative were then asked how arousing they found those feelings. Arousal ratings from participants in the positive condition (when rating the felt arousal in connection with experienced positive affect) and the negative condition (when rating the arousal felt in connection with experienced negative affect) were not significantly different, $t(211.11) = 0.18$, $p = 0.86$, $d = 0.02$ (see Fig. 2).

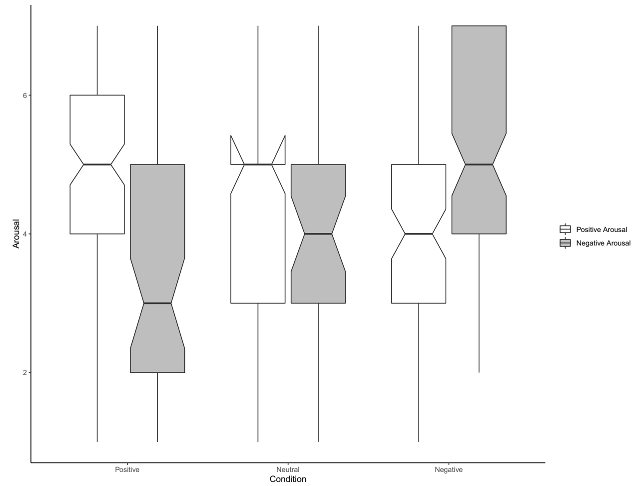

Fig. 2 Positive and negative arousal of speech task by condition. The two outside boxes display similar levels of arousal for participants in the positive and negative conditions. Boxes show values between the 25% and 75% quantiles; dark horizontal lines represent median value; indented "notches" represent 95% confidence interval; whiskers represent range of responses

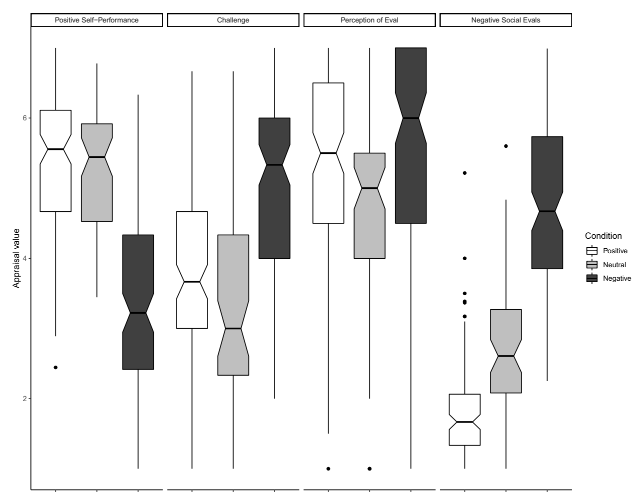

Fig. 3 Speech performance appraisals by condition. Boxes show values between the 25% and 75% quantiles; dark horizontal lines represent median value; indented "notches" represent 95% confidence interval; whiskers represent range of non-outlying values; dots represent outlying values

Appraisals

All four ANOVAs comparing task appraisals as a function of condition were significant. See Fig. 3 for a graphical depiction and Table 2 for full results. Planned contrasts revealed that participants in the positive and neutral conditions had similar levels of positive self-performance ratings, which were higher than ratings from those in the negative condition. The negative condition had higher

Table 2 Mean performance appraisals by condition

	Positive condition		Neutral condition		Negative condition		Omnibus
Appraisal	*M* (*SE*)	95% CI	*M* (*SE*)	95% CI	*M* (*SE*)	95% CI	*F* (2, 298)
Positive self-performance	$5.34 (0.09)_a$	[5.16, 5.53]	$5.26 (0.11)_a$	[5.04, 5.48]	$3.38 (0.12)_b$	[3.14, 3.61]	111.20
Challenge	$3.73 (0.13)_a$	[3.48, 3.97]	$3.39 (0.16)_a$	[3.06, 3.71]	$5.08 (0.12)_b$	[4.84, 5.32]	44.70
Perception of evaluation	$5.29 (0.13)_a$	[5.04, 5.55]	$4.62 (0.18)_b$	[4.26, 4.97]	$5.61 (0.12)_a$	[5.38, 5.84]	11.34
Negative social evaluations	$1.81 (0.07)_a$	[1.68, 1.94]	$2.75 (0.12)_b$	[2.51, 2.99]	$4.79 (0.11)_c$	[4.57, 5.01]	273.43

Note: means with different subscripts differ from each other at $p \leq 0.001$; all omnibus tests significant at $p < 0.001$. Sample sizes of each condition: positive ($n = 117$), neutral ($n = 64$), and negative ($n = 120$)

Table 3 Measures of physiological arousal by condition and time

		Baseline		Speech-task		Recovery	
Measure	Condition	*M* (*SE*)	95% CI	*M* (*SE*)	95% CI	*M* (*SE*)	95% CI
Mean arterial pressure	Positive	75.52 (0.78)	[73.98, 77.07]	99.14 (1.17)	[96.83, 101.45]	80.07 (0.89)	[78.34, 81.80]
	Neutral	77.54 (1.35)	[74.88, 80.19]	98.24 (2.01)	[94.27, 102.22]	80.92 (1.51)	[77.94, 83.90]
	Negative	75.84 (0.77)	[74.32, 77.35]	99.75 (1.15)	[97.49, 102.02]	81.57 (0.86)	[79.87, 83.27]
Heart rate	Positive	75.10 (1.27)	[72.59, 77.60]	90.44 (1.82)	[86.85, 94.03]	76.37 (1.27)	[73.88, 78.87]
	Neutral	73.06 (2.19)	[68.75, 77.38]	89.04 (3.13)	[82.86, 95.22]	74.00 (2.18)	[69.70, 78.29]
	Negative	72.00 (1.25)	[69.54, 74.46]	94.41 (1.79)	[90.88, 97.93]	74.43 (1.24)	[71.98, 76.88]

Note: sample sizes of each condition: positive ($n = 80$), neutral ($n = 27$), and negative ($n = 83$)

challenge ratings than the positive condition, which in turn had higher scores on this composite than the neutral condition. This same pattern was observed in perception of evaluation ratings, although the difference between the negative and positive groups was not statistically significant ($p = 0.074$). Finally, the negative group reported more negative social evaluations, followed by the neutral group, and the positive group.

Physiological Arousal

MAP and HR followed a similar trajectory for participants in each condition (see Table 3). Each first increased significantly from baseline to during the speech and then decreased significantly from the speech to recovery, main effects of time; MAP: $F(1.55, 289.21) = 899.12$, $p < 0.001$, $\eta_p^2 = 0.83$; HR: $F(1.27, 236.64) = 275.43$, $p < 0.001$, $\eta_p^2 = 0.60$. For MAP, there was no time by condition interaction, $F(3.09, 289.21) = 1.49$, $p = 0.22$, $\eta_p^2 = 0.02$, illustrating that participants in all three conditions showed similar trajectories on MAP over the course of the session. For HR, however, a significant interaction between time and condition, $F(2.53, 236.64) = 6.10$, $p = 0.001$, $\eta_p^2 = 0.06$, indicated that participants in the negative condition had a greater increase in HR from baseline to speech compared to the positive and neutral conditions, but recovered to have a similar HR to the other two conditions after the speech had finished.

Discussion

The TSST is one of the most potent and widely used laboratory stressors, helping researchers understand how people function in a stressful situation by providing a controlled laboratory induction to assess the effects of emotion on physiological and psychological outcomes. However, to date, no non-stressful positive versions of the TSST have been able to reliably increase positive affect from a baseline state to a greater extent than other TSST conditions while also equating subjective emotional arousal. The present research developed a successful positive TSST procedure, where the level of post-task positive affect differed significantly across all three conditions in the intended direction. Using the present paradigm, the positive TSST increased positive affect and decreased negative affect in participants; the neutral TSST did not change levels of positive affect and led to a decrease in negative affect; and the negative TSST increased negative and decreased positive affect.

Importantly, participants in the positive condition (when rating arousal of the positive affect) and participants in the negative condition (when rating the arousal of the negative affect) did not differ in reported levels of arousal. Thus, the manipulation was successful in isolating an effect of valence and not arousal and demonstrated that people can perceive non-stressful positive experiences to be just as arousing as negative experiences. This self-reported finding of arousal was corroborated by the BP measurements taken throughout the session. The speech conditions demonstrated similar trajectories for MAP and HR across the session, although HR increased significantly more during the TSST for those in the negative condition. These initial results suggest similar physiological profiles between conditions, but that the negative condition may be slightly more physiologically arousing. The finding of elevated HR in association with the negative condition during the TSST is consistent with previous research showing that, beyond initial HR reactivity to emotional states (which is similar for both positive and negative affect; Jacob et al., 1999), HR remains elevated longer in response to negative compared to positive affect (Brosschot & Thayer, 2003). Of note, however, BP and HR are influenced by both the sympathetic and parasympathetic nervous systems (Cacioppo et al., 2000; Mauss & Robinson, 2009). Therefore, we caution against drawing any strong conclusions with regards to physiological arousal using only these metrics and encourage future research to include additional measures that reflect only sympathetic activity (e.g., pre-ejection period) or only parasympathetic activity (e.g., heart rate variability).

Past research using a positive TSST activated the sympathetic nervous system, which, combined with measures of increased cardiac efficiency, indicated participants were experiencing a stress response characteristic of "challenging" (as opposed to "threatening") situations (Blascovich & Mendes, 2000). Our positive manipulation is unique in that it is arousing (as evidenced by BP ratings and self-reported arousal), yet not considered stressful, challenging, or threatening by participants. Therefore, the initial evidence supports that the present manipulation provides a non-stressful, positive, yet arousing comparison condition to the classic TSST.

We introduced several small differences between conditions in an effort to make a successful positive TSST condition. Although we believe all of these changes together made our manipulation more successful, one limitation of this approach is that we are unable to isolate exactly what part of the manipulation makes the positive TSST successful. Furthermore, differences between conditions could have inadvertently altered other aspects of the task related to valence, such as approach/avoid motivations and goal relevance. What drives the success of the positive manipulation, as well as what other aspects related to valence may be influenced, could be investigated in future research. We believe it is the combination of a supportive audience with an enjoyable speech topic that makes this iteration of a positive TSST successful. To diminish differences between conditions, while still inducing positive or negative emotion in participants, we suggest researchers attempt having an enjoyable speech topic in all conditions and only vary the behavior/responses of the evaluators.

The present research employed a sample of undergraduate students, which limits the generalizability of our results. As with all existing iterations of the TSST, the validity of the manipulations should be examined in different age and cultural groups to ensure generalizability. This is particularly important given that culture influences the type of positive emotions individuals value most (high arousal or low arousal; Tsai et al., 2006). Given that our sample was comprised of a relatively ethnically diverse group of individuals, however, the findings presented here provide initial evidence that the present methodology would be widely suitable to a diverse American undergraduate sample.

The manipulations presented in this research successfully produced experimental conditions that differed in post-speech positive affect. For negative affect, the negative condition did increase negative affect more so than the other two conditions. Yet, there were no post-speech differences in negative affect between the neutral and positive conditions. This appears to be due to the neutral speech task decreasing negative affect among participants in the neutral condition, despite our goal to maintain levels of negative affect in this condition.

Given that the classic TSST procedure produces high-arousal negative affect, we designed our positive TSST equivalent to yield high-arousal positive affect in order to differentiate conditions based on valence but not arousal. As a result, research using the manipulations presented here is limited to investigations of high-arousal affective states and should not be generalized to situations involving low-arousal affective states. Future paradigms could potentially be developed to compare the effects of valence in low-arousal situations. However, high-arousal situations (e.g., the "peak" or most intense portions of an emotional experience) often have the greatest impact on psychological processes (Fredrickson, 2000), thus, we expect these manipulations to be useful for investigating several aspects of psychological phenomena. Some potential applications of this positive, non-stressful TSST equivalent could include tests of how valence differentially influences attention, memory accuracy, decision making, or creativity. Further, researchers could test how carefully matched positive and negative social experiences influence emotional responses as a function of individual difference characteristics, such as social anhedonia or having a history of depression.

Taken together, the research presented here demonstrates the validity of a new high-arousal positive, yet non-stressful laboratory procedure, one that can easily be compared to the pre-existing and widespread negative TSST speech task. It is our hope that future research can utilize this ecologically valid paradigm to answer questions about how positive and negative affect differ in their effects on a variety of psychological and physiological outcomes.

Acknowledgements The authors would like to thank researchers Cailah, Candace, Celeste, Cesar, Christina, Colleen, Erick, Erika, Gao, Hannah, Jasmine, Jasveen, Jennifer, Joanna, Johanna, Kate, Lauren, Leslie, Mary, Megan, Missy, Mohini, Monika, Paige, Paula, Phoebe, Saum, Stella, Susan, Xueping, and Zhanna for their efforts during protocol development and data collection, and for making this research possible.

Additional Information

Funding E.J.U. was supported by National Institute of Mental Health T32MH018931. The research was supported by a dissertation research award from the American Psychological Association.

Data Availability Study materials, computed variables, and analytic syntax are available on the Open Science Framework at: https://osf.io/hkvqs/?view_only=6336c6421cd8471aa855909efb1fa7d0.

Ethical Approval The procedure presented in the study was designed as part of the first author's dissertation. All procedures were approved by the University's Institutional Review Board.

Conflict of Interests The authors declare no competing interests.

Informed Consent Informed consent was obtained from all research participants.

Disclaimer The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42761-021-00074-6.

References

Akinola, M., & Mendes, W. B. (2008). The dark side of creativity: Biological vulnerability and negative emotions lead to greater artistic creativity. *Personality and Social Psychology Bulletin, 34*(12), 1677–1686. https://doi.org/10.1177/0146167208323933

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323–370. https://doi.org/10.1037//1089-2680.5.4.323

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: the role of affective cues. In J. P. Forgas (Ed.), *Feeling and thinking The role of affect in social cognition* (pp. 59–82). Cambridge University Press.

Brosschot, J. F., & Thayer, J. F. (2003). Heart rate response is longer after negative emotions than after positive emotions. *International Journal of Psychophysiology, 50*(3), 181–187. https://doi.org/10.1016/S0167-8760(03)00146-6

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. Haviland-Jones (Eds.), *The Handbook of emotions* (pp. 173–191). Guildford Press.

Champely, S. (2018). pwr: Basic functions for power analysis. R package version 1.2–2. https://CRAN.R-project.org/package=pwr.

Compton, R. J. (2003). The interface between emotion and attention: a review of evidence from psychology and neuroscience. *Behavioral and Cognitive Neuroscience Reviews, 2*, 115–129. https://doi.org/10.1177/1534582303002002003

Crum, A. J., Akinola, M., Martin, A., & Fath, S. (2017). The role of stress mindset in shaping cognitive, emotional, and physiological responses to challenging and threatening stress. *Anxiety, Stress and Coping, 30*(4), 379–395. https://doi.org/10.1080/10615806.2016.1275585

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Dickerson, S. S., Mycek, P. J., & Zaldivar, F. (2008). Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association, 27*(1), 116–121. https://doi.org/10.1037/0278-6133.27.1.116

Dickerson, S. S., Gable, S. L., Irwin, M. R., Aziz, N., & Kemeny, M. E. (2009). Social-evaluative threat and proinflammatory cytokine regulation. *Psychological Science, 20*(10), 1237–1244. https://doi.org/10.1111/j.1467-9280.2009.02437.x

Dolcos, F., Labar, K. S., & Cabeza, R. (2004). Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: An event-related fMRI study. *NeuroImage, 23*(1), 64–74. https://doi.org/10.1016/j.neuroimage.2004.05.015

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition & Emotion, 14*(4), 577–606. https://doi.org/10.1080/026999300402808

Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology, 74*(5), 1337–1349. https://doi.org/10.1037/0022-3514.74.5.1337

Herten, N., Otto, T., Adolph, D., Pause, B. M., Kumsta, R., & Wolf, O. T. (2016). Enhanced startle responsivity 24 hours after acute stress exposure. *Behavioral Neuroscience*, *130*(4). https://doi.org/10.1037/bne0000156.

Howell, R. T., Kern, M. L., & Lyubomirsky, S. (2007). Health benefits: Meta-analytically determining the impact of well-being on objective health outcomes. *Health Psychology Review, 1*(1), 83–136. https://doi.org/10.1080/17437190701492486

Jacob, R. G., Thayer, J. F., Manuck, S. B., Muldoon, M. F., Tamres, L. K., Williams, D. M., Ding, Y., & Gatsonis, C. (1999). Ambulatory blood pressure responses and the circumplex model of mood: A 4-day study. *Psychosomatic Medicine, 61*(3), 319–333. https://doi.org/10.1097/00006842-199905000-00011

Kassam, K. S., Koslov, K., & Mendes, W. B. (2009). Decisions under distress: Stress profiles influence anchoring and adjustment. *Psychological Science, 20*(11), 1394–1399. https://doi.org/10.1111/j.1467-9280.2009.02455.x

Kassam, K. S., & Mendes, W. B. (2013). The effects of measuring emotion: Physiological reactions to emotional situations depend on whether someone is asking. *PLoS ONE, 8*(6), 6–13. https://doi.org/10.1371/journal.pone.0064959

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review, 1*(2), 99–113. https://doi.org/10.1177/1754073908100432

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective & Behavioral Neuroscience, 6*(2), 110–126. https://doi.org/10.3758/CABN.6.2.110

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.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66*, 799–823. https://doi.org/10.1146/annurev-psych-010213-115043

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A Grumpy overview. *Social Cognition, 22*(5), 530–554.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131*(6), 803–855. https://doi.org/10.1037/0033-2909.131.6.803

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion, 23*(2), 209–237. https://doi.org/10.1080/02699930802204677.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *Manual for the Profile of Mood States (POMS)*. Educational and Industrial Testing Service.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231–259.

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin, 125*(1), 3–30. https://doi.org/10.1037//0033-2909.125.1.3

Sin, N. L., Moskowitz, J. T., & Whooley, M. A. (2015). Positive affect and health behaviors across 5 years in patients with coronary heart disease: The heart and soul study. *Psychosomatic Medicine, 77*(9), 1058–1066. https://doi.org/10.1097/PSY.0000000000000238

Talarico, J. M., Berntsen, D., & Rubin, D. C. (2009). Positive emotions enhance recall of peripheral details. *Cognition and Emotion, 23*(2), 380–398. https://doi.org/10.1080/02699930801993999

Taylor, S. E., Seeman, T. E., Eisenberger, N. I., Kozanian, T. A., Moore, A. N., & Moons, W. G. (2010). Effects of a supportive or an unsupportive audience on biological and psychological responses to stress. *Journal of Personality and Social Psychology, 98*(1), 47–56. https://doi.org/10.1037/a0016563

Torre, J. B., & Lieberman, M. D. (2018). Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review, 10*(2), 116–124. https://doi.org/10.1177/1754073917742706

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology, 90*(2), 288–307. https://doi.org/10.1037/0022-3514.90.2.288

Waters, S. F., West, T. V., & Mendes, W. B. (2014). Stress contagion: Physiological covariation between mothers and infants. *Psychological Science, 25*(4), 934–942. https://doi.org/10.1177/0956797613518352

Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. http://www.ncbi.nlm.nih.gov/pubmed/3397865.

Weidman, A. C., Steckler, C. M., & Tracy, J. L. (2016). The jingle and jangle of emotion assessment: Imprecise measurement, casual scale usage, and conceptual fuzziness in emotion research. *Emotion*. https://doi.org/10.1037/emo0000226

Wiemers, U. S., Sauvage, M. M., Schoofs, D., Hamacher-Dang, T. C., & Wolf, O. T. (2013a). What we remember from a stressful episode. *Psychoneuroendocrinology, 38*, 2268–2277. https://doi.org/10.1016/j.psyneuen.2013.04.015

Wiemers, U. S., Sauvage, M. M., & Wolf, O. T. (2013b). Odors as effective retrieval cues for stressful episodes. *Neurobiology of Learning and Memory, 112*, 230–236. https://doi.org/10.1016/j.nlm.2013.10.004

Wiemers, U. S., Schoofs, D., & Wolf, O. T. (2013c). A friendly version of the trier social stress test does not activate the HPA axis in healthy men and women. *Stress (Amsterdam, Netherlands), 16*(2), 254–260. https://doi.org/10.3109/10253890.2012.714427

Wiemers, U. S., & Wolf, O. T. (2015). Cortisol broadens memory of a non-stressful social interaction. *Psychopharmacology (berl), 232*(10), 1727–1733. https://doi.org/10.1007/s00213-014-3808-7