Relationship of Cortisol and Alpha Amylase to Behavioral Engagement under Three Levels of Negative Evaluative Psychosocial Stress

By: Alessandra R. Grillo, Gail M. Courneau, Suzanne Vrshek-Schallhorn

Grillo, A.R., Corneau, G.M., Vrshek-Schallhorn, S. Relationship of Cortisol and Alpha Amylase to Behavioral Engagement under Three Levels of Negative Evaluative Psychosocial Stress. *Psychoneuroendocrinology, 155*, 106308.

Abstract:

Despite that behavioral engagement is integral to mental health, surprisingly little is known about the relationship of psychosocial stress and behavioral engagement. The current study developed an observer-rated measure of behavioral engagement for lab-based stress inductions, then examined its relationship with stress-responsive biomarkers and affect. Young adults (N = 109, Mage=19.4, SDage=1.59, 57% female) completed one of three Trier Social Stress Test (TSST) conditions—non-stressful Control, Intermediate, or an Explicit Negative Evaluative—and at four timepoints provided self-reports of positive and negative affect and saliva samples for cortisol and salivary alpha-amylase (sAA). Trained study staff (experimenters and TSST judges) completed a programmed questionnaire measure of the novel behavioral engagement measure after the participants completed the TSST. Psychometric review and EFA of the behavioral engagement items resulted in a final 8-item measure with good interrater reliability and well-fitting 2-factor structure, capturing Persistence (4 items; loadings=.41–.89), and Quality of Speech (4 items; loadings=.53–.92). Results indicated that the relationship of positive affect growth and biomarker level to behavioral engagement varied substantially as a function of context: As negative evaluation level strengthened, behavioral engagement became more tightly associated with relative preservation of positive affect. For both cortisol and sAA, the relationship between biomarker levels (but not reactivity) and behavioral engagement varied significantly by condition, such that under milder conditions and elevated levels of biomarkers, engagement was greater, but under Explicit Negative Evaluation, and elevated levels of biomarkers, engagement was less, suggesting behavioral withdrawal. Findings reveal the critical role of context—especially negative evaluation—in the relationship of biomarkers with behavioral engagement.

Keywords: Negative evaluation | Trier social stress test | Behavioral engagement | Cortisol | Alpha-amylase | Affect | Scale development

Article:

1. Introduction

Since its inception in the 1990 s, the Trier Social Stress Test (Kirschbaum et al., 1993) has markedly advanced our understanding of biological, cognitive, and affective responses to psychosocial threat. Absent from this body of work, however, is much understanding of behavioral responses to the TSST. While this gap is remarkable because reduced behavioral engagement is a correlate of stress-related conditions such as depression (Horne et al., 2021, Treadway et al., 2009), it may not be surprising given that there is no psychometric measure of behavioral engagement during the TSST. Further, the relations of behavioral engagement and stress-responsive biomarkers—and how these relationships vary as a function of context, the level of psychosocial threat—are unknown. To address these gaps, the current study: 1) developed and evaluated the psychometric properties of a measure of behavioral engagement for the TSST (Behavioral Indicators of TSST Engagement, "BITE" scale) and 2) examined behavioral engagement in relation to both the level and reactivity in affect, cortisol, and salivary alpha-amylase (sAA) during three levels of a TSST paradigm in emerging adults.

1.1. Relevance of Behavioral Engagement under Psychosocial Stress.

Behavioral engagement is a multidisciplinary construct defined in slightly different ways across several literatures (Fredricks et al., 2004, Hughes et al., 2008, O'Keefe et al., 2017). In the education literature, effortful engagement refers to the volitional aspect of involvement in, and attentional resources to, instructional (e.g., classroom) activities, and includes persistence in the face of difficulty (Hughes et al., 2008). Definitions of engagement within the psychological literature vary, but broadly coalesce on engagement as the observable involvement in an activity, and as a downstream outcome of motivation (see Martin et al., 2017 for a discussion). Additional theory and evidence suggest that levels of effort during a task depend on the person's level of interest in and appraised importance of the task, perceived attainability of success, and critical to the current project, whether the task is self-relevant (Brehm and Self, 1989, Gendolla and Richter, 2010, O'Keefe et al., 2017).

Self-relevance in the context of task engagement is defined as settings in which people's performance has consequences and implications for their self-definition and self-esteem (Gendolla and Richter, 2010). There are multiple theoretical accounts of how self-relevance facilitates motivation, such as social determination theory (Deci and Ryan, 2000) which posits that the need for relatedness, competence, and autonomy from others can spur motivation. Moreover, the social self-preservation model (Dickerson et al., 2004) shows how acute stress relates to this process. This theory posits that social evaluation mobilizes a set of coordinated psychological and biological responses to maintain and protect one's self-image when under perceived acute threat. These two theories show self-relevance as a catalyst for motivation, however, neither have yet been extended to account for downstream behavioral engagement. Because past research conceptualizes motivation as "the impetus to" engagement (Martin et al., 2017), it follows that in contexts when a task is self-relevant, and there is threat to one's self-image (e.g., during psychosocial stress), the perceived threat engenders motivation, which ought to influence behavioral engagement. The acute stressor used in the current study, the Trier Social Stress Test (Kirschbaum et al., 1993), is thought to elicit an acute stress response because of its negative evaluative component, which can threaten the self, thereby influencing behavioral engagement.

Behavioral engagement is worthwhile to examine for two reasons. First, levels of behavioral engagement—or disengagement—are correlates of stress-related disorders like depression (Treadway et al., 2009). For example, people with elevated depressive symptoms demonstrated blunted cardiac activity—a marker of behavioral withdrawal—as difficulty on a cognitive task increased (Silvia et al., 2016). Further, people with higher depressive anhedonia exerted less effort to obtain rewards compared to those with lower anhedonia (Treadway et al., 2009). Yet, while behavioral engagement may be relevant to internalizing risk, no measure has assessed behavioral engagement during the TSST. Second, observational measures of behavioral engagement are likely to offer a complement to self-reported perceived effort due to evidence of bias in performance self-appraisals (Scheiter et al., 2020). Therefore, a measure of behavioral engagement during the TSST may provide an accurate external assessment of effort during task performances.

1.2. Importance of negative evaluation level in stress responding

Given the importance of social self-preservation to motivation, the type of TSST used to examine behavioral engagement is relevant. Recent investigations suggest that the level of negative evaluation in lab-based stress inductions is critical to responses (Woody et al., 2018). In contrast to the original TSST in which judges behave coolly and neutrally (ambiguously negative evaluative), a recent variation has judges deliver explicit (both verbal and non-verbal) negative evaluative feedback (Way and Taylor, 2010; described later in methods). This variant produced significantly more cortisol reactivity on average than a variant like the original TSST (Vrshek-Schallhorn et al., 2018). Importantly, however, a depression risk factor (trait rumination) predicted higher cortisol reactivity to the variant like the original TSST, but blunted cortisol reactivity in an explicit negative evaluative variant (Vrshek-Schallhorn et al., 2018). The latter finding also replicated in another sample (Vrshek-Schallhorn et al., 2019). In conceptualizing cortisol as a resource-mobilizing hormone (Sapolsky et al., 2000), blunted cortisol reactivity under negative evaluative threat could be indicative of a "giving up" physiological response (Vrshek-Schallhorn et al., 2018). We therefore hypothesized that, on average, explicit negative evaluation might inhibit behavioral engagement, whereas ambiguous evaluation might facilitate behavioral engagement compared to a non-stressful Control (i.e., a non-linear effect of negative evaluative level on behavioral engagement). Thus, the present study sought to develop a measure to use with three levels of evaluation.

1.3. Relationships of behavioral engagement to affect under stress

Prior work demonstrates a link between behavioral engagement and positive affect (extent to which a person feels enthusiastic, active, and alert; Watson et al., 1988), during task performance (O'Keefe et al., 2017). While, in another study with three TSST conditions, negative affect (aversive mood states, such as anger, contempt, disgust, guilt; Watson et al., 1988) reactivity increased approximately linearly as negative evaluative level strengthened (Vrshek-Schallhorn et al., 2018). To examine associations of behavioral engagement with affect across three evaluative levels during the TSST, we hypothesized that greater behavioral engagement would positively correlate with PA reactivity and would negatively correlate with NA reactivity, and that this relationship would intensify as negative evaluative level strengthened, i.e., an interaction with condition.

1.4. Relationships of behavioral engagement to biomarkers under stress

Predicting the relationship between behavioral engagement and both levels and reactivity in cortisol and salivary alpha-amylase (sAA, an indicator of sympathetic reactivity under stimulated conditions; Granger and Taylor, 2020; Nater and Rohleder, 2009) proves to be more complex. Baseline levels of each biomarker immediately prior to the TSST may causally influence subsequent behavioral engagement because each serves a preparatory function to respond to threat (Sapolsky et al., 2000), but elevated levels of each biomarker have been linked to both behavioral inhibition (e.g., freezing responses, Roelofs, 2017; Sherman and Mehta, 2020) and to behavioral activation, such as fight or flight responses or motivated performances (Dickerson and Kemeny, 2004, Nater et al., 2006).

Regarding sAA, theory and evidence contend that energy mobilization for task engagement is mediated by cardiovascular activity, particularly beta-adrenergic activity (e.g., see Richter et al., 2008), and sAA is a marker of adrenergic activity (Rohleder et al., 2004). Specifically, it is thought that beta-adrenergic activity increases with task difficulty to a certain point, but once the task is appraised as too difficult, beta-adrenergic levels decrease accordingly (Richter et al., 2008). Regarding cortisol responses, the "boost" hypothesis posits that elevations in cortisol will predict greater subsequent activeness and alertness, both central to engagement (Hoyt et al., 2016). Complicating matters, the magnitude of cortisol response (augmented or blunted) that should be considered "adaptive" appears to be context-dependent, specifically on negative evaluation level: an internalizing risk indicator predicted greater cortisol reactivity to the ambiguous Intermediate condition of the TSST, a possible psychophysiological "overreaction," but also predicted blunted cortisol reactivity to the Explicit Negative Evaluative condition (Vrshek-Schallhorn et al., 2018). This suggests that the association between behavioral engagement and cortisol reactivity could vary, even non-linearly, across levels of negative evaluation. Thus, although we hypothesized that behavioral engagement would associate with biomarker levels and reactivity, and that these relationships would vary by condition (potentially non-linearly for cortisol), because of previously reported mixed findings, we did not specify the direction of expected effects.

1.5. The present study

This study evaluated a new measure of behavioral engagement designed for the TSST and examined its relationship with affective and physiological units of analyses. We developed an observer-rated behavioral engagement questionnaire, the "BITE" scale, completed by TSST experimenters and judges, and examined its psychometric properties across three objective levels of evaluation, a non-evaluative Control, an ambiguous Intermediate, and an Explicit Negative Evaluative condition. We hypothesized that behavioral engagement would be greatest in the Intermediate condition (Hypothesis 1). We then hypothesized that behavioral engagement would associate with changes in PA (Hypothesis 2) and NA (Hypothesis 3) over time as a function of condition. Next, to probe relationships between behavioral engagement and biomarkers, in separate growth curve models for cortisol and sAA, we simultaneously tested non-directional hypotheses that behavioral engagement would be associated with level (Hypotheses 4 and 5) and reactivity (Hypothesis 6 and 7) over time as a function of condition, respectively.

2. Method

2.1. Participants

Participants at a Southeastern U.S. university were recruited through psychology courses and completed an eligibility questionnaire. Exclusion criteria were current use of nicotine, hormonal birth control, psychotropic or corticosteroid medications, chronic health conditions, and several criteria pertaining to cognitive tests for aims unrelated to the current analyses (learning disability, uncorrected vision, hearing deficits, a first language other than English, and history of head trauma). Eligibility was also contingent on a healthy blood pressure screening for aims unrelated to the current study.

Following consent, participants (N = 145) were screened for active depression using the Structured Clinical Interview for DSM-IV (First et al., 2015). If participants met criteria for depression, they were switched to the Control condition due to associations of depression with cortisol reactivity (Burke et al., 2005), and were excluded from analyses due to non-randomization (n = 6). We further excluded participants for outlying cortisol levels (n = 4) in preliminary inspection; no participants had outlying sAA levels. We also excluded participants who withdrew prior to completion (n = 8). For consistency across conditions, only the experimenters' ratings (and not judges' ratings) were used in the primary analyses. In 30 cases, however, the experimenter did not to complete the BITE questionnaire immediately following the session and participants were excluded from analyses. Of the 109 participants in the analytic sample, demographic information was available for n = 108. Participants identified as 38.0% White; 33.3% Black or African American; 12.0% Biracial; 7.4% as Asian or Pacific Islander; 3.7% as Hispanic or Latine; 0.9% as American Indian or Alaskan Native, and 4.6% as Other. Participants were 18–29 years (M = 19.4, SD = 1.59), and 57% identified as female, and 43% male. Participants had on average, 13.7 years of education.

2.2. Procedure

All procedures were approved by the university's Institutional Review Board and all participants provided informed consent. Study sessions took place in the afternoon to avoid confounding with the diurnal rhythm of cortisol and sAA (Nater et al., 2006). Participants were quasi-randomized (i.e., signed up for study appointments unaware of prescheduled conditions) to complete one condition of the TSST: the putatively non-stressful Control condition (n = 42), the moderately stressful Intermediate condition (n = 41), or the robustly Explicit Negative Evaluative condition (n = 26).

2.2.1. TSST protocol

All conditions of the TSST included five minutes of speech task preparation, five minutes of speaking on a pre-assigned topic, and five minutes of mental arithmetic. Participants were permitted to make notes during the preparation period but were not permitted to use them during the speech task. During the arithmetic portion, participants were instructed to sequentially subtract 13 from 2017 as quickly and accurately as possible. If participants made a mistake, they were required to start over from the beginning. In all three conditions, participants completed the speech

and arithmetic tasks facing a video camera and were told they were being video recorded. In reality, no recordings were made.

The Control, Intermediate, and Explicit Negative Evaluative conditions differed in four ways to increase stressor severity. First, in the Control and Intermediate conditions, participants were explicitly informed that their performance would not be evaluated, while in the Explicit Negative Evaluative condition, participants were told that they would be evaluated both by the judges via questionnaires and later by experts through analysis of their video recordings. Second, in the Control condition, no confederate judges were present, although the experimenter remained in the room. This person sat out of the participant's direct eyeshot and was trained to appear minimally attentive to the participant's performance. In the Intermediate and Explicit Negative Evaluative conditions, one and two judges were present, respectively. To avoid confounding with demographic variables of judges and participants, the two Explicit Negative Evaluative condition judges differed from each other with respect to both gender identity (one male identifying, one female) and race/ethnicity. Third, the evaluative demeanor of the judges differed across conditions. In the Intermediate condition, the judge made eye contact with the participant and appeared engaged but neutral in response to the participant's performance; in the Explicit Negative Evaluative condition, the two judges made eye contact with the participant and provided negative verbal and non-verbal feedback according to a behavioral script (one confederate was trained to appear bored and the other to appear dissatisfied). Fourth, the level of evaluation inherent in the speech topic strengthened across the three conditions, from a non-evaluative topic in the Control condition (tips for living a healthy lifestyle), to a mildly evaluative topic in the Intermediate condition (actions the participant would take as a leader in a student organization), to an explicitly self-evaluative topic in the Explicit Negative Evaluative condition (reasons a participant would be the best person to elect to a leadership position).

After participants completed the TSST protocol, the experimenters and judges (depending on condition) completed the BITE measure to assess the participant's observed behavioral engagement during the TSST, for a total of one rating per participant in the Control and two ratings in the Intermediate and Explicit Negative Evaluative conditions. In the Explicit Negative Evaluative condition, the two judges completed one BITE through consensus discussion.

2.3. Measures

2.3.1. Socioeconomic status (SES)

Participants completed the Hollingshead Index (Hollingshead, 1975) to provide socioeconomic data on their family of origin.

2.3.2. Salivary cortisol and alpha-amylase

Participants provided samples of saliva via passive drool into sterile cryogenic vials at four time points: (1) at baseline (+0 min), (2) after instructions and the TSST including the preparation period (+30 min after baseline), (3) after cognitive tasks not presented here (+65 min after baseline), and (4) after debriefing and further relaxation (+75 min after baseline). Vials were stored at -80 °C and shipped to The University of Trier Biochemisches Lab in Trier, Germany. Salivary cortisol was duplicate assayed using time-resolved fluorescent-detection immunoassay (DELFIA; Dressendörfer et al., 1992), while sAA was duplicate assayed using a 2-chloro-4-nitrophenyl-

alpha-D-maltotrioside method and spectrophotometric measurement (Lorentz et al., 1999). Cortisol intra-assay coefficient of variation (CV) ranged from 4.0% to 6.7% and inter-assay variation ranged from 7.0% to 9.0%. Further, sAA intra-assay CV ranged from 2.8% to 6.3% and inter-assay CV ranged from 5.5% to 7.6%. All four timepoints were used for cortisol analyses to capture reactivity; the first three were used for sAA analyses because average levels fully recovered by the third time point as expected, then slightly increased again at the fourth timepoint following sAA's diurnal rhythm (Nater et al., 2007).

2.3.3. Self-reported affect

Participants completed the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) at the same four timepoints as saliva collection to assess momentary PA and NA. The PANAS is sensitive to change in affect and is recommended for use during the TSST (Narvaez Linares et al., 2020). Analyses included the first three time points (capturing before and after the TSST and recovery) because any reactivity in PA and NA returned to baseline by the third timepoint, similar to sAA. The 20-item PANAS includes 10 items assessing PA and NA respectively on a 5-point Likert scale (1-very slightly or not at all to 5-extremely), which are analyzed as separate subscales. Internal consistency for each subscale ranged from acceptable to good across four timepoints (PA subscale α = .847–.908, and NA subscale α = .598–.758).

2.3.4. Novel measure of behavioral engagement

Scale development of the "BITE" measure involved drafting of an initial item pool by two researchers with experience with the TSST, followed by a feedback phase from three experienced researchers and a group of research assistants experienced in serving as judges and experimenters. We conceptualized behavioral engagement on the TSST as a spectrum of observable behaviors ranging from failing to invest in the task (either due to not taking the task seriously or to being excessively inhibited) to appearing fully invested (taking the task seriously and confidently engaging). We expected that maximum behavioral engagement would include: 1) fully using the preparation period, 2) staying on task for the full duration of the speech and mathematic portions, and 3) requiring few reminders from judges to stay on task. Specific to the speech task, we posited that maximum engagement would include: 1) delivering an organized speech, 2) incorporating novel content throughout the speech, 3) providing specific details rather than generalizations, and 4) following instructions to use the speech topic provided rather than making negative self-relevant comments ("I'm doing a bad job"). Specific to the math task, we anticipated that individuals making a greater effort would exhibit fewer errors and achieve a better (i.e., lower) final number on the arithmetic task. Five items had intentionally very similar content, administered once tailored to the speech portion and once tailored to the math portion. We anticipated that a speech and a math subscale would emerge.

The initial pool had 19 items: one item regarding the participants' use of the preparation period, 11 regarding engagement and performance during the speech, and 7 regarding engagement and performance during the math portion. We began with a larger pool than necessary, retaining items with strong psychometric performance for use in a brief final measure for experimenters to complete. Items were developed with some redundancy—the same construct was assessed in opposite directions, and the same constructs were assessed separately for the speech and math portions. Items reflecting apparent defeated behavior or not taking the task seriously were reverse

scored, so that in the final scoring, 0 reflected low engagement and 4 reflected high engagement (see Table S7 in Supplement for original measure). Study staff were trained by the PI (SVS) to rate the measure prior to its administration and completed it as a Qualtrics questionnaire immediately after the participant completed the study session.

2.4. Analytic plan

2.4.1. Preliminary analyses

First, we screened data for excessive outlying biomarker values (>M+/−3 SDs) in cortisol and sAA in the combined sample for baseline samples, and within condition for remaining samples, using Area Under the Curve with Respect to Increase (AUCi; Pruessner et al., 2003) values, and winsorized outliers to M+/− 3 SDs, which is customary (Vrshek-Schallhorn et al., 2018). If AUCi values still exceeded M+/− 3 SDs, they were excluded from all analyses (n = 4 for cortisol, n = 0 for sAA). Cortisol reactivity scores were somewhat skewed (statistic = 1.496) and kurtotic (statistic = 3.158), as were sAA reactivity values (skew = 1.962, kurtosis = 5.294), warranting natural log transformation of both biomarkers, consistent with conventions (Byrne, 2016). Next, we conducted one-way ANOVAs to rule out differences by condition in demographic variables. Final preliminary analyses included running simple growth curve models to demonstrate whether biomarker and affect levels were impacted as expected by increasing levels of negative evaluation, reflected as an interaction of Stress Condition x Quadratic Time for all variables, using SAS Version 9.4 for growth curve modeling (SAS Institute, 2013). We used alpha levels of $p < .05$ for all analyses.

2.4.2. Psychometric properties of the BITE

To assess psychometric properties of the BITE, we used experimenter ratings for all but inter-rater reliabilities so that all conditions would be treated equally given the number of judges varied by condition. First, we examined item-level descriptive statistics, including item distribution and skew in the whole sample (Worthington and Whittaker, 2006). Items demonstrating excessive skew or other non-normality across conditions were removed (initial items 8, 11, 12, 13, 16, 17, 18 and 19). Second, we examined item-scale correlations on the BITE measure; item correlations <.3 were removed per best practices (Watkins, 2018), resulting in removal of initial item 1 ($r = .29$). Distributions were then further inspected within condition, resulting in no additional exclusions. Third, scale-level analyses included exploratory factor analysis (EFA) to examine dimensionality of the remaining items and multilevel growth curve modeling with the subscales indicated by the EFA to examine the predictive validity of the BITE with cortisol, sAA, PA, and NA reactivity during the TSST. We used SPSS Version 26 (IBM Corp, 2019) for descriptive and correlational analyses, MPlus Version 8 for EFA (Muthén and Muthén, 2017). Indices to assess model fit include 1) root mean square error of approximation (RMSEA), where values >.05 indicate close fit, between .05 and .08 indicate reasonable fit, and values surpassing .10 indicate poor fit (Brown and Cudek, 1993), 2) standardized root mean square residual (SRMR), with values >.10 suggesting a poor fit, and 3) the comparative fit index (CFI) and Tucker-Lewis Index (TLI), where values >.95 indicate good fit (Kline, 2015). Next, we computed intraclass correlations to examine interrater-reliability of remaining items between the confederate judge or judges' rating and the experimenter's rating (Shrout and Fleiss, 1979), with values >.5 indicative of moderate reliability,

>.7 good reliability, and >.9 indicating excellent reliability (Koo and Li, 2016). Data were analyzed as item-means in growth curve models to permit missing up to 1 item.

Multilevel growth curve analysis used the item-mean score of the BITE (which we refer to as "total" to indicate the full scale rather than subscales), but because EFA identified a two-factor solution, supplemental results (Tables S1-S6) are presented for the two subscales. To test the relation of stress condition on behavioral engagement (Hypothesis 1), we examined a one-way ANOVA with polynomial contrasts.

2.4.3. Growth curve analyses

Multilevel growth models examined BITE item-mean score for the full scale (and, in supplemental results, item-means for the two subscales) in relation to three elements of repeatedly measured biomarkers: intercept (overall level, influenced by anticipatory effects at baseline), linear effect of time (upward or downward growth from baseline), and quadratic effect of time (upward and downward trajectory, or curvilinearity, reflecting reactivity from baseline). Due to prior evidence of curvilinear effects of negative evaluative condition on cortisol (Vrshek-Schallhorn et al., 2018), models for cortisol incorporated both linear and quadratic (i.e., curvilinear) effects of condition. These multilevel models allow for examination of nested data (i.e., timepoints at level 1 nested within people at level 2). The BITE item-mean score was grand-mean centered in all models. The time variables used orthogonal coefficients to achieve uncorrelated terms for linear and quadratic effects of time. To account for uneven spacing in sample measurements (e.g., measurements collected at 0, +30, +65, and +75 min), time was coded as 0, 2, 4, and 5 before creating orthogonal coefficients of time for cortisol, but even spacing was acceptable for sAA, NA, and PA, each modeled at three timepoints (collected at 0, +30, and +65 min). Covariance was unstructured. Preliminary analyses indicated a random intercept-slope model for biomarkers and a random intercept model for affect variables.

First, for each outcome (PA, NA, cortisol, sAA), we provide simplified growth curve models, using only the effects of Condition, Linear Time, and Quadratic Time to most clearly convey the effects of stress condition on the linear growth and quadratic reactivity of each.

Second, we add the BITE to models for each outcome to test the relation of behavioral engagement and PA (Hypothesis 2) and NA (Hypothesis 3) we examined both linear and quadratic growth of affect across condition. Specifically, we examined a three-way interaction between BITE x Linear Time x Condition and BITE x Quadratic Time x Condition to assess the link between behavioral engagement and NA and PA reactivity, respectively. Post-hoc decomposition for affect models entailed examining two-way interactions of BITE x Linear Time and BITE x Quadratic Time for each condition using simple slope analyses to assess associations between engagement and affect growth as stressor severity strengthened.

To test our hypotheses of cortisol level (Hypothesis 4) and its relation to behavioral engagement across condition, we examined the BITE x Condition effect (intercept effect), accounting for both linear and quadratic effects of Condition (BITE x Condition, BITE x Condition2) due to prior evidence of non-linear effects of condition with psychological variables, described above. A BITE x Condition intercept effect would reflect that levels of biomarker are associated with BITE score as a function of condition. For sAA level (Hypothesis 5), analyses were the same except we only accounted for the linear effect of condition in the absence of prior evidence for non-linear effects of condition. For cortisol reactivity (Hypothesis 6) with behavioral engagement, we tested if there are differences in change over time as a function of condition,

reflected in interactions with linear growth or curvilinear reactivity (e.g., BITE x Condition x Quadratic Time). To examine sAA reactivity with behavioral engagement (Hypothesis 7), we examined BITE x Condition x Quadratic Time, testing only linear effects of condition. Across all models, post-hoc probing of significant interaction effects used simple slope analyses for each condition.

3. Results

3.1. Preliminary analyses

One-way ANOVAs revealed no differences between conditions on age, SES, baseline levels of cortisol or sAA, or baseline NA and PA (ps =.19–.80). Chi-square tests showed no differences between conditions on gender or minority status (ps = .68–.99). One-way ANOVAs further showed no differences in BITE scores by self-reported gender across (p = .71) or within (ps = .06–.71) condition, or by minority status across (p = .93) or within (ps = .49–.78) condition (Table 1).

Table 1. Descriptive Values by Condition.

	Control (n = 42)	**Intermediate (n = 41)**	**Explicit Negative Evaluative (n = 26)**
Self-Report Gender	M = 23, F = 19	M = 24, F = 17	M = 12, F = 14
	M (SD)	M (SD)	M (SD)
Age	19.52 (1.69)	19.00 (1.21)	19.58 (1.90)
Baseline Positive Affect (Item mean)	2.44 (.80)	2.63 (.89)	2.48 (.55)
Baseline Negative Affect (Item mean)	1.30 (.41)	1.28 (.35)	1.24 (.27)
Baseline Cortisol	3.34 (1.72)	3.82 (2.41)	3.53 (1.89)
Baseline Salivary Alpha-Amylase	4.48 (.82)	4.77 (.85)	4.54 (.64)
BITE Total	22.5 (6.83)	20.79 (6.77)	20.00 (7.07)

Note: Total refers here to the item-mean full measure. Results for subscales appear in the supplement.

3.2. Exploratory factor analysis (EFA)

After initial item examinations and interrater reliability calculations, 10 items remained for the EFA. We tested 10 items using maximum-likelihood estimation with oblimin (geomin) rotation in MPlus Version 8, because we anticipated correlated factors (Muthén and Muthén, 2017). Items with poor loadings were dropped iteratively one at a time, rerunning the EFA with each item removal (Worthington and Whittaker, 2006), resulting in deletion of two items (items 7 and 15). Model fit indices indicated that a 2-factor model achieved good fit (RMSEA=.019, CFI=.999, TLI=.997, SRMR=.039). Four items loaded onto Factor 1 (coefficients = .41–.89), and four items loaded onto Factor 2 (coefficients = .53–.92; Table 2).

Table 2. Exploratory Factor Analysis Results for the Behavioral Indicators of Trier Engagement (BITE) Measure.

Scale Items	Factor Loadings (λ)	
	Persistence	**Quality of Speech**
I estimate that the participant spoke for: 0 – less than half of the 5 min 1 – about two-thirds of the 5 min 2 – about three-quarters of the 5 min 3 – almost but not quite the full 5 min 4 – for the full 5 min	0.815	0.117
How often did the participant need reminders to continue speaking? (R) 0 – almost always 1 – frequently 2 – sometimes 3 – rarely 4 – never	0.888	-0.009
*During the speech portion, the participant appeared defeated (R)	0.519	0.282
*During the math portion, the participant appeared defeated (R)	0.406	-0.003
*The participant's speech content was delivered in an organized fashion. For example, the participant transitioned from point to point smoothly, rather than in a manner that was difficult to follow	0.003	0.919
*The participant's speech content was repetitive (R)	-0.214	0.797
*The participant provided details in his/or her speech, rather than superficial generalizations	0.142	0.525
*During the speech, the participant appeared motivated and engaged	0.059	0.616
Factor Eigenvalues	4.0871	1.114

Note. All factor coefficients (λ ≥ .40) are formatted in bold font.
*Items experimenters/confederates completed items using this scale: 0 – strongly disagree, 1 – somewhat disagree, 2 – neither agree nor disagree, 3 – somewhat agree, 4 – strongly agree
(R) Item was reverse scored

Contrary to our initial predictions, the factors did not reflect performance for the speech and math portions separately, due in part to the poor psychometric performance and subsequent deletion of the math items. Instead, Factor 1 appeared to represent the extent of observed Persistence, measured by the amount of time spoken (item 2), need for reminders (item 3), and maintenance of composure during both the speech (item 9) and math (item 14) tasks, for a total of 3 speech items and 1 math item. Factor 2 appeared to index the Quality of Speech, reflecting the level organization, novel content, and details in the speech (items 4, 5, and 6, respectively), as well as the participant's apparent motivation during the speech (item 10). As expected, the two factors were significantly correlated ($r = .67$, $p < .05$).

Interrater reliability yielded acceptable intraclass correlations as follows: .783 for the total BITE scale, .835 for the Persistence subscale, and .712 for the Quality of Speech subscale.

3.3. Effect of negative evaluation level on behavioral engagement

To examine the effect of negative evaluation level on behavioral engagement, a one-way ANOVA with polynomial contrasts was conducted. We predicted behavioral engagement would be highest

in the Intermediate condition (Hypothesis 1). Contrary to predictions, results indicated that behavioral engagement did not vary significantly as either a quadratic ($F = .119$, $p = .731$) or linear ($F = 2.132$, $p = .147$) function of negative evaluation level. Instead, behavioral engagement was similar across conditions. Subscale analyses revealed similar patterns.

3.4. Multilevel growth curve models

3.4.1. Effect of stress condition on all outcomes

We first ran simple models to examine the effect of TSST condition on reactivity in all outcomes, that is, models including only Condition, Linear Time, Quadratic Time, and their interaction effects to predict PA, NA, sAA, and cortisol respectively. The manipulation significantly altered reactivity for all outcomes, given by Condition x Quadratic Time interactions (Table 3). Specifically, for PA, the Condition x Quadratic Time effect ($t = 3.71$, $p = .0003$) indicated a change from modest positive PA reactivity in the Control to general declines in PA over time in more intense conditions. By contrast, for NA, cortisol, and sAA, as condition strengthened, reactivity increased (interaction effect ts from −6.49 to −3.64, ps from <.0001 to.0003; more negative effect sizes indicate larger inverted U-curves across repeated measures, i.e., reactivity). Because patterns of reactivity roughly corresponded to expectations, we proceeded with growth curves examining the relationship of the BITE to each outcome.

3.4.2. Positive and negative affect associations with behavioral engagement

To test the relation of behavioral engagement and PA (Hypothesis 2) and NA (Hypothesis 3), we examined both linear and quadratic growth of affect across condition. Specifically, we examined a three-way interaction between BITE x Linear Time x Condition and BITE x Quadratic Time x Condition to assess NA and PA, respectively (Table 4, Table 5). The relationship between PA and behavioral engagement varied significantly by condition, a BITE x Linear Time x Condition interaction ($\beta = 0.0105$, $SE(\beta) = 0.0045$, $t(234) = 2.30$, $p = .0223$; see Table 4). Post-hoc simple slope analyses for each condition showed that in the Control condition, more behavioral engagement was not significantly (but negatively) associated with growth in positive affect over time ($\beta = -0.00896$, $SE(\beta = 0.005876$, $t(234) = -1.52$, $p = .129$), whereas in the Intermediate condition, more behavioral engagement was not significantly associated with growth in PA over time ($\beta = 0.001495$, $SE(\beta) = 0.003705$, $t(234) = 0.40$, $p = .687$), and in the Explicit Negative Evaluative condition, greater behavioral engagement was significantly associated with growth in PA over time ($\beta = 0.0120$, $SE(\beta) = 0.0059$, $t(234) = 2.04$, $p = .0423$). Specifically, under the Explicit Negative Evaluative condition, people with greater behavioral engagement experienced relatively less decline in self-reported PA (Fig. 1). Results for behavioral engagement total score and subscales of Persistence and Quality of Speech were similar (Supplement S1 and S4). We did not detect a relationship between behavioral engagement and PA quadratic growth (BITE x Quadratic Time x Condition ($\beta = 0.000135$, $SE(\beta) = 0.002624$, $t(234) = 0.05$, $p = .9591$). Results indicate that the relationship of behavioral engagement to PA varies significantly by condition, and that greater behavioral engagement was significantly associated with relative preservation of PA in the Explicit Negative Evaluative condition only.

Table 3. Simple Effects of Condition on all Outcomes.

A. Positive Affect Model	**b**	**SE (b)**	**DF**	**t-value**	**p-value**
Intercept	2.3746	0.1056	120	22.5	<.0001****
Linear Time	-0.208	0.03983	240	-5.22	<.0001****
Quadratic Time	-0.09321	0.023	240	-4.05	<.0001****
Condition	-0.1094	0.08414	240	-1.3	0.1949
Condition x Linear Time	-0.04893	0.03175	240	-1.54	0.1246
Condition x Quadratic Time	0.06798	0.01833	240	3.71	0.0003***
B. Negative Affect Model	**b**	**SE(b)**	**DF**	**t-value**	**p-value**
Intercept	1.2751	0.0523	120	24.38	<.0001****
Linear Time	-0.08225	0.03737	240	-2.2	0.0287*
Quadratic Time	-0.06171	0.02158	240	-2.86	0.0046**
Condition	0.1352	0.04169	240	3.24	0.0013**
Condition x Linear Time	0.05359	0.02979	240	1.8	0.0733
Condition x Quadratic Time	-0.1116	0.0172	240	-6.49	<.0001****
C. Cortisol Model	**b**	**SE(b)**	**DF**	**t-value**	**p-value**
Intercept	0.9579	0.07388	248	12.96	<.0001****
Linear Time	-0.2845	0.06385	123	-4.46	<.0001****
Quadratic Time	-0.08399	0.03488	248	-2.41	0.0168*
Condition	0.212	0.05961	248	3.56	0.0004**
Condition x Linear Time	0.1813	0.05152	248	3.52	0.0005**
Condition x Quadratic Time	-0.126	0.02815	248	-4.48	<.0001****
D. sAA Model	**b**	**SE(b)**	**DF**	**t-value**	**p-value**
Intercept	4.5964	0.1081	123	42.52	<.0001****
Linear Time	-0.02428	0.0278	246	-0.87	0.3834
Quadratic Time	-0.05531	0.01605	246	-3.45	0.0007***
Condition	0.06787	0.08722	246	0.78	0.4373
Condition x Linear Time	0.000714	0.02243	246	0.03	0.9746
Condition x Quadratic Time	-0.04715	0.01295	246	-3.64	0.0003***

*p < .05; **p < .01; ***p < .001, ****p < .0001

Table 4. Growth Curve Results for BITE, Condition, and Time with Positive Affect.

	b	SE (b)	DF	t-value	p-value
Intercept	2.3536	0.1083	117	21.72	<.0001***
Time	-0.2020	0.04063	234	-4.97	<.0001***
Time2	-0.08574	0.02346	234	-3.66	0.0003***
Condition	-0.08262	0.08572	234	-0.96	0.3361
BITE Total	0.01470	0.01567	234	0.94	0.3492
Condition x Time	-0.04509	0.03214	234	-1.40	0.1620
Condition x Time2	0.06066	0.01856	234	3.27	0.0012**
BITE Total x Time	-0.00896	0.005876	234	-1.52	0.1288
BITE Total x Time2	-0.00426	0.003393	234	-1.25	0.2108
BITE Total x Condition	0.002894	0.01212	234	0.24	0.8115
BITE Total x Time x Condition	0.01045	0.004545	234	2.30	0.0223*
BITE Total x Time2 x Condition	0.000135	0.002624	234	0.05	0.9591

Note. BITE Total was centered.
*p < .05; **p < .01; ***p < .001

Table 5. Growth Curve Results for BITE, Condition, and Time with Negative Affect.

	b	SE (b)	DF	t-value	p-value
Intercept	1.2766	0.05434	117	23.49	<.0001***
Time	-0.09407	0.03870	234	-2.43	0.0158*
Time2	-0.06402	0.02235	234	-2.86	0.0046**
Condition	0.1314	0.04300	234	3.06	0.0025**
BITE Total	0.001171	0.007861	234	0.15	0.8818
Condition x Time	0.06027	0.03062	234	1.97	0.0502
Condition x Time2	-0.1092	0.01768	234	-6.18	<.0001***
BITE Total x Time	0.007893	0.005599	234	1.41	0.1599
BITE Total x Time2	0.001297	0.003232	234	0.40	0.6886
BITE Total x Condition	-0.00285	0.006080	234	-0.47	0.6399
BITE Total x Time x Condition	-0.00457	0.004330	234	-1.06	0.2924
BITE Total x Time2 x Condition	0.000101	0.002500	234	0.04	0.9679

Note. BITE Total was centered.
*p < .05; **p < .01; ***p < .001

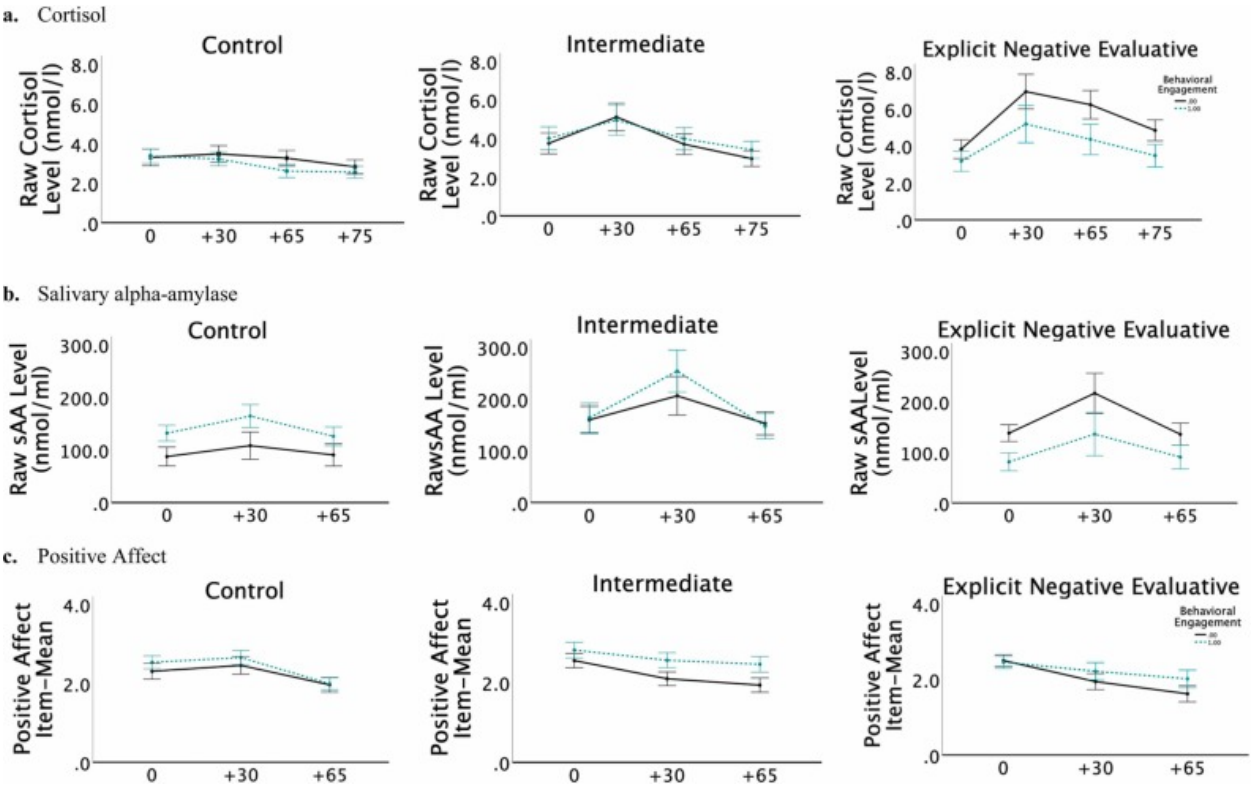

Fig. 1. Biomarker and Affective Outcomes by Stress Condition, Behavioral Engagement Score, and Time. Note. Behavioral Engagement Score is Median Split (Low= 4 – 22.99; High= 23 – 28.00); Bars represent +/−1 standard error.

Regarding Hypotheses 3, we detected no relationship of behavioral engagement to NA growth or reactivity as a function of condition (BITE x Linear Time x Condition ($\beta = 0.007893$, $SE(\beta) = 0.005599$, $t(234) = 1.41$, $p = .1599$; BITE x Quadratic Time x Condition ($\beta = 0.001297$, $SE(\beta) = 0.003232$, $t(234) = 0.40$, $p = .6886$); see Table 5.

3.4.3. Cortisol level, growth, and reactivity relationship to behavioral engagement

Moreover, the relationship of cortisol level with behavioral engagement varied significantly as a function of condition, coded both linearly (BITE x Condition: $\beta = 0.08735$, $SE(\beta) = 0.02940$, $t(236) = 2.97$, $p = .0032$) and quadratically (BITE x Condition2: $\beta = -0.04713$, $SE(\beta) = 0.01414$, $t(236) = -3.33$, $p = .0010$), Hypothesis 4. By contrast, regarding Hypothesis 6, there were no significant relationships between cortisol growth or reactivity and behavioral engagement, whether modeling condition linearly or quadratically, and whether examining linear growth or quadratic reactivity; (e.g., no significant three-way interactions; ps from.112 to.192; Table 6).

To characterize the findings for behavioral engagement and cortisol level, we examined simple slopes by condition. In the Control condition, cortisol level was negatively but not significantly associated with behavioral engagement ($\beta = -0.01629$, $SE(\beta) = 0.01166$, $t(236) = -1.40$, $p = .1634$), whereas in the Intermediate condition, cortisol level was positively and significantly associated with behavioral engagement ($\beta = 0.02393$, $SE(\beta) = 0.01144$, $t(354) = 2.09$, $p = .0371$) and finally in the Explicit Negative Evaluative condition, cortisol level was negatively and significantly associated with behavioral engagement ($\beta = -0.03010$, $SE(\beta) = 0.01183$, $t(354) = -2.54$, $p = .0114$). As such, cortisol level was associated with facilitation of behavioral

Table 6. Growth Curve Results for BITE, Condition, and Time with Cortisol.

	b	SE (b)	DF	t-value	p-value
Intercept	0.9813	0.08125	236	12.08	<.0001***
Time	-0.1901	0.07280	117	-2.61	0.0102*
Time2	-0.09260	0.04073	236	-2.27	0.0239*
Condition	0.1273	0.1994	236	0.64	0.5238
Condition2	0.02987	0.09611	236	0.31	0.7562
BITE Total	-0.01629	0.01166	236	-1.40	0.1638
Condition x Time	-0.1725	0.1787	236	-0.97	0.3354
Condition x Time2	-0.1203	0.09997	236	-1.20	0.2302
Condition2 x Time	0.1649	0.08611	236	1.91	0.0567
Condition2 x Time2	-0.00070	0.04817	236	-0.01	0.9884
BITE Total x Time	-0.00802	0.01045	236	-0.77	0.4437
BITE Total x Time2	0.009179	0.005847	236	1.57	0.1178
BITE Total x Condition	0.08735	0.02940	236	2.97	0.0033**
BITE Total x Condition2	-0.04713	0.01414	236	-3.33	0.0010**
BITE Total x Condition x Time	0.03919	0.02634	236	1.49	0.1381
BITE Total x Condition x Time2	-0.02352	0.01474	236	-1.60	0.1118
BITE Total x Condition2 x Time	-0.01940	0.01266	236	-1.53	0.1269
BITE Total x Condition2 x Time2	0.009281	0.007085	236	1.31	0.1915

Note. BITE Total was centered.
*p < .05; **p < .01; ***p < .001

To characterize the findings for behavioral engagement and cortisol level, we examined simple slopes by condition. In the Control condition, cortisol level was negatively but not significantly associated with behavioral engagement ($\beta = -0.01629$, $SE(\beta) = 0.01166$, $t(236) = -1.40$, $p = .1634$), whereas in the Intermediate condition, cortisol level was positively and significantly associated with behavioral engagement ($\beta = 0.02393$, $SE(\beta) = 0.01144$, $t(354) = 2.09$, $p = .0371$) and finally in the Explicit Negative Evaluative condition, cortisol level was negatively and significantly associated with behavioral engagement ($\beta = -0.03010$, $SE(\beta) = 0.01183$, $t(354) = -2.54$, $p = .0114$). As such, cortisol level was associated with facilitation of behavioral engagement under the milder ambiguous negative evaluation in the Intermediate condition but inhibition of behavioral engagement under explicit negative evaluation in the Explicit Negative Evaluative condition (Fig. 1). Effects were similar across both BITE subscales (Supplement Tables S2 and S5).

3.4.4. sAA level, growth, and reactivity relationship to behavioral engagement

The model indicated that the relationship of sAA level with behavioral engagement varied significantly as a function of condition, a BITE x Condition interaction effect, $\beta = -0.03850$, $SE(\beta) = 0.01226$, $t(236) = -3.14$, $p = .0019$, suggesting that as negative evaluative level of the condition strengthened, the association between sAA level and behavioral engagement became more negative (Hypothesis 5). By contrast, there were no significant interaction effects of BITE score with linear time ($\beta = 0.000947$, $SE(\beta) = 0.001858$, $t(236) = 0.51$, $p = .611$) or quadratic time ($\beta = -0.00061$, $SE(\beta) = 0.003217$, $t(236) = -0.19$, $p = .850$), indicating no significant relationship between sAA growth or reactivity during the TSST and behavioral engagement (Hypothesis 7).

The model further showed a significant main effect such that BITE score predicted sAA ($\beta = -0.04515$, SE(β)= 0.01573, $t(236) = 2.87$, $p = .0045$), see Table 7.

Table 7. Growth Curve Results for BITE, Condition, and Time with Salivary Alpha-Amylase.

	b	SE (b)	DF	t-value	p-value
Intercept	4.5475	0.1090	118	41.71	<.0001***
Time	-0.02767	0.02860	236	-0.97	0.3344
Time2	-0.05505	0.01651	236	-3.33	0.0010**
Condition	0.07843	0.08686	236	0.90	0.3674
BITE Total	0.04515	0.01573	236	2.87	0.0045**
Condition x Time	0.002866	0.02279	236	0.13	0.9000
Condition x Time2	-0.04910	0.01316	236	-3.73	0.0002**
BITE Total x Time	0.001267	0.004126	236	0.31	0.7591
BITE Total x Time2	-0.00444	0.002382	236	-1.86	0.0637
BITE Total x Condition	-0.03850	0.01226	236	-3.14	0.0019**
BITE Total x Time x Condition	-0.00061	0.003217	236	-0.19	0.8496
BITE Total x Time2 x Condition	0.000947	0.001858	236	0.51	0.6105

Note. BITE Total was centered.
*p < .05; **p < .01; ***p < .001

Simple slope analyses to decompose the BITE x Condition interaction indicated that in the Control condition, sAA level was significantly and positively associated with behavioral engagement ($\beta = 0.04515$, SE(β)= 0.01573, $t(236) = 2.87$, $p = .0045$). In the Intermediate condition, there was no significant association between sAA level and behavioral engagement ($\beta = 0.006647$, SE(β)= 0.01003, $t(236) = 0.66$, $p = .5080$). In the Explicit Negative Evaluative condition, however, sAA level was significantly and negatively associated with behavioral engagement ($\beta = -0.03185$, SE(β)= 0.01595, $t(236) = -2.00$, $p = .0470$). In parallel to findings for cortisol, sAA level was associated with facilitation of behavioral engagement under only the mildest putatively non-stressful Control condition, but inhibition of behavioral engagement in the Explicit Negative-Evaluative condition. Similar patterns were observed for both the Persistence and Quality of Speech subscales (Supplement Tables S3 and S6).

4. Discussion

The current study provided initial validation of an observer-rated psychometric measure of behavioral engagement during a lab-based psychosocial stress task, and for the first time, showed that the relationship of positive affect growth and biomarker level to behavioral engagement varies substantially as a function of context. Specifically, for both cortisol and alpha-amylase, the relationship between biomarker levels (but not reactivity) and behavioral engagement varied significantly by condition, such that under milder conditions (i.e., Control for sAA but Intermediate for cortisol), elevated biomarkers were associated with greater engagement, but under Explicit Negative Evaluation, they were associated with less engagement. Further, results demonstrated that changes in PA (but not NA) were associated with behavioral engagement dependent on condition: People with greater behavioral engagement had less decline in self-reported positive affect under Explicit Negative-Evaluation but not under conditions with little to moderate evaluative threat where average PA declines were less pronounced.

4.1. Initial validation of a behavioral engagement measure

The present report provides evidence of initial validity for the first psychometric measure developed to measure behavioral engagement during lab-based stress inductions, the Behavioral Indicators of TSST Engagement (BITE) scale. Exploratory factor analysis identified a two-factor solution for the BITE—Persistence and Quality of Speech—with very good fit; internal consistency reliability for both the two subscales and an overall score were also both good. Further, interrater reliability measured in two conditions between the experimenter and one or more judges was also strong, leading us to conclude that the BITE is a psychometrically sound measure useful for elements of future lab-based stress induction research.

4.2. Biomarker relationships with behavioral engagement

Biomarkers levels were associated with behavioral engagement as a function of environmental conditions, such that higher sAA level was associated with more behavioral engagement in the putatively non-stressful condition, higher cortisol level was associated with more behavioral engagement in the Intermediate condition, and higher levels of both biomarkers were associated with reduced engagement in the high stress Explicit Negative Evaluative condition. Both sAA and cortisol level may have even causally influenced behavioral engagement, given that baseline measures contributing to intercept effects indeed occurred prior to the behavior (and moreover, represent even earlier levels due to time for biomarkers to migrate into saliva). However, because these results are in part cross-sectional (deriving in part from biomarker measures occurring simultaneous with behavior), firm causal interpretations await experimental research externally manipulating cortisol and sAA level pharmacologically and observing behavioral engagement during TSST performance. Contrary to predictions, however, the current study did not support that either cortisol or sAA (linear) growth or (quadratic) reactivity related significantly to behavioral engagement during the TSST over and above contributions of baseline biomarker levels, which suggests that anticipatory levels may be especially important for behavior measured over brief time courses such as in the TSST. While previous research identifies increases in biomarkers in anticipation of forthcoming challenge (MacDonald and Wetherell, 2019), to the best of our knowledge, only one study has examined the predictive value of baseline sympathetic activity on subsequent task performance or levels of engagement in the context of a challenge. Matsumura et al. (2021) found that sympathetic activity, indicated by sympathetic tone, prior to engaging in competitive skating predicted a stronger performance during the competition compared to non-competitive practice sessions. By contrast, the present results suggest that under the most intense interpersonal pressure, that high levels of biomarkers were associated with poorer engagement, i.e., behavioral withdrawal. Motivational intensity theory (Brehm and Self, 1989) posits that differences in appraisals of a task's difficulty can impact motivation and effort. The competitive skaters in Matsumura et al. (2021) likely knew the level of difficulty of the task and gave commensurate effort. The relative unexpectedness of the task in the current study likely precluded accurate appraisals of difficulty, which in turn, may have affected effort and engagement.

Further, these patterns of both cortisol and sAA activity on engagement could align in part with evidence regarding challenge and threat states (Blascovich & Mendes, 2000), although these constructs have predominantly been examined relative to sympathetic functioning and not HPA axis functioning. Specifically, a challenge state occurs when people perceive the attainability of

the situation, such that their personal coping resources match or exceed situational demands, which leads to sharpened executive functioning mediated by sympathetic activity. By contrast, a threat state occurs when people perceive that the situation is too demanding and they do not possess adequate resources, which has in prior reports correlated with reduced sympathetic activity and, in turn, dampened decision making and subsequent performance (Hase et al., 2019). In the current study, biomarkers were measured in the context that the participant had agreed to complete a challenging task protocol with speech and math components, but were also aware that they did not have full information about what it would entail. Participants' stress responsive systems could have been stimulated by the ambiguous, impending challenging task, mobilizing resources accordingly. Participants in the Control, non-stressful condition likely perceived their ability to complete the tasks and therefore benefited from the enhancing effects of sympathetic arousal, as was the case for HPA activation in the Intermediate condition, influencing task engagement. By contrast, the Explicit Negative Evaluative condition likely provoked feelings of unattainability where not enough sympathetic arousal was available to prepare for the challenge at hand, leading to poorer engagement (e.g., less persistence in giving speech or following instructions). Importantly, the relationship of findings to "challenge" and "threat" states is speculative, the current study was not a direct test of this theory, and the current results diverge from the "threat" states (Blascovich & Mendes, 2000), such that reduced sympathetic activity was not observed in the current study. However, this points to important directions in future research that might examine continuous sympathetic activity throughout a psychosocial challenge to gain a granular understanding of changes in sympathetic activity prior to the onset of a challenge and while the challenge is occurring.

Similarly, cortisol has been conceptualized as a resource-mobilizing hormone that initiates physiological processes that function to prepare a person to act (Sapolsky et al., 2000) or compensates for a deficit in other resources such as sleep or poor mood, in accord with the boost hypothesis (Hoyt et al., 2016). Additionally, evidence from studies using TSST has long suggested that greater cortisol reactivity in some circumstances is a maladaptive pattern of stress responding (Vreeburg et al., 2009), and emerging work interestingly points to cortisol's causal role in adverse responding. It is thought that heightened cortisol influences avoidance or freezing behaviors in the context of stressful stimuli (Roelofs et al., 2005). Specifically, participants exposed to the TSST who had high cortisol responses then had decreased approach-avoidance behavior (Roelofs et al., 2005). Taken together across studies, cortisol may function as a facilitator or inhibitor on behavior depending on environmental context. In line with this logic, the results of the current study demonstrate cortisol as a resource mobilizing hormone that facilitates behavioral engagement in the context of mild pressure, but inhibits behavior under maximally negative evaluative pressure. The present evidence may thus help reconcile seemingly contrasting views of cortisol.

4.3. Relationship of behavioral engagement to affect

The current study also demonstrated that context was a significant factor in understanding the relationship between behavioral engagement and PA. Under lower levels of stress and negative evaluation, no relationship was detected, but in the Explicit Negative Evaluative condition, people with greater behavioral engagement had less decline in PA. The construct of PA reflects elements such as activeness, determination, inspiration, and interest in the environment (Wróbel et al., 2019). Although the present data provide a cross-sectional snapshot of behavior and self-reported affect, in clinical contexts, stimulating behavioral activation has been shown to causally bolster

positive emotion and reduce anhedonia (Dimidjian et al., 2011); future research could apply experimental designs manipulating level of behavioral engagement under stress to assess its impact on PA. But conversely, it is also possible that relatively persisting PA facilitated behavioral engagement under stress, given the known role of PA in stimulating coping (Khosla, 2006), or that the two mutually influenced one another.

By contrast, although NA behaved as expected across increasing levels of negative evaluation, there were no significant associations between NA level or reactivity and behavioral engagement in the current study. Although a rich body of work implicates NA in behavioral inhibition (Polivy, 1998), and the BITE measure was intended to capture a spectrum of withdrawn inhibition to active engagement, NA has also been notoriously difficult to correlate to other indicators of stress reactivity in the TSST, so much that a qualitative review failed to link NA reactivity to biomarker reactivity at all (Campbell and Ehlert, 2012).

4.4. Limitations

Although the current study developed and incorporated a novel observed-rated measure of behavioral engagement tailored to a well-validated, frequently used psychosocial stress paradigm, and used multiple physiologically indicators of stress reactivity, and employed three conditions, it is not without limitations. First, increased saliva and affect sampling frequency (e.g., every 5–10 min) may provide cleaner resolution on momentary processes. Second, candidate items on the BITE pertaining to math task engagement tended to perform poorly psychometrically and many were removed from the final version of the measure; future studies could seek to revise the BITE and investigate other types of questions pertaining to math task engagement. Third, the current study's three conditions of the TSST spanned the spectrum of negative evaluation but not explicit positive evaluation, which has been shown to produce sizable cortisol reactivity (Way and Taylor, 2010). Positive evaluations are likely relevant to behavioral engagement and may be an important direction for future research. Fourth, the current study did not measure IQ, which could be a potential indicator of external validity for the Quality of Speech subscale. Last, it is possible that there could indeed be relationships between biomarker reactivity and behavioral engagement that our study was underpowered to detect.

5. Conclusion

The current study provides initial validation for a novel measure of behavioral engagement during lab-based stress inductions and provides evidence that levels of salivary biomarkers are associated with subsequent behavioral engagement as a function of context, linking elevated biomarkers to greater behavioral engagement under milder pressure but inhibited engagement under more pressure, and highlighting the importance of future research using an Explicit Negative Evaluative TSST condition. This work will facilitate the study of predictors and sequelae of behavior under stressful conditions and supports the critical role of environmental interpersonal context in modulating both boosting and inhibiting functions of sympathetic activation and cortisol.

Declaration of Competing Interest

The authors Alessandra Grillo, Gail Corneau, and Suzanne Vrshek-Schallhorn have no known conflict of interests to disclose.

References

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

Brehm, J.W., Self, E.A., 1989. The intensity of motivation. Annu. Rev. Psychol. 40, 109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545.

Brown, M., Cudek, R., 1993. Alternative ways of assesing model fit in: Bollen.

Burke, H.M., Davis, M.C., Otte, C., Mohr, D.C., 2005. Depression and cortisol responses to psychological stress: a meta-analysis. Psychoneuroendocrinology 30, 846–856. https://doi.org/10.1016/j.psyneuen.2005.02.010.

Byrne, B.M., 2016. Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming, Third Edition. Routledge.

Campbell, J., Ehlert, U., 2012. Acute psychosocial stress: does the emotional stress response correspond with physiological responses? Psychoneuroendocrinology 37, 1111–1134. https://doi.org/10.1016/j.psyneuen.2011.12.010.

Deci, E.L., Ryan, R.M., 2000. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol. Inq. 11, 227–268. https://doi.org/10.1207/S15327965PLI1104_01.

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.

Dickerson, S.S., Gruenewald, T.L., Kemeny, M.E., 2004. When the social self is threatened: shame, physiology, and health. J. Pers. 72, 1191–1216. https://doi.org/ 10.1111/j.1467-6494.2004.00295.x.

Dimidjian, S., Barrera, M., Martell, C., Mu˜noz, R.F., Lewinsohn, P.M., 2011. The origins and current status of behavioral activation treatments for depression. Annu. Rev. Clin. Psychol. 7, 1–38. https://doi.org/10.1146/annurev-clinpsy-032210-104535.

Dressend¨orfer, R.A., Kirschbaum, C., Rohde, W., Stahl, F., Strasburger, C.J., 1992. Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. J. Steroid Biochem. Mol. Biol. 43, 683–692. https://doi.org/10.1016/0960-0760(92)90294-S.

First, M.B., Williams Janet, B., Karg, R., Spitzer, R., 2015. Structured Clinical Interview for DSM-5 Research Version. Am. Psychiatr. Assoc.

Fredricks, J.A., Blumenfeld, P.C., Paris, A.H., 2004. School engagement: potential of the concept, state of the evidence. Rev. Educ. Res. 74, 59–109. https://doi.org/10.3102/00346543074001059.

Gendolla, G.H.E., Richter, M., 2010. Effort mobilization when the self is involved: some lessons from the cardiovascular system. Rev. Gen. Psychol. 14, 212–226. https://doi.org/10.1037/a0019742.

Granger, D.A., Taylor, M.K., 2020. Salivary Bioscience Foundations of Interdisciplinary Saliva Research and Applications.

Hase, A., O'Brien, J., Moore, L.J., Freeman, P., 2019. The relationship between challenge and threat states and performance: a systematic review. Sport Exerc. Perform. Psychol. 8, 123–144. https://doi.org/10.1037/spy0000132.

Hollingshead, A.B., 1975. Four Factor Index of Social Status [WWW Document]. URL https://sociology.yale.edu/sites/default/files/files/yjs_fall_2011.pdf#page=21 (Accessed 4.26.20).

Horne, S.J., Topp, T.E., Quigley, L., 2021. Depression and the willingness to expend cognitive and physical effort for rewards: a systematic review. Clin. Psychol. Rev. 88, 102065 https://doi.org/10.1016/j.cpr.2021.102065.

Hoyt, L.T., Zeiders, K.H., Ehrlich, K.B., Adam, E.K., 2016. Positive upshots of cortisol in everyday life. Emotion 16, 431–435. https://doi.org/10.1037/emo0000174.

Hughes, J.N., Luo, W., Kwok, O.-M., Loyd, L.K., 2008. Teacher-student support, effortful engagement, and achievement: a 3-year longitudinal study. J. Educ. Psychol. 100, 1–14. https://doi.org/10.1037/0022-0663.100.1.1.

IBM Corp, 2019. SPSS v26. Khosla, M., 2006. Positive Affect and Coping with Stress 9.

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.

Kline, R.B., 2015. Principles and Practice of Structural Equation Modeling, Fourth ed.,. Guilford Publications.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropr. Med. 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Lorentz, K., Gütschow, B., Renner, F., 1999. Evaluation of a direct alpha-amylase assay using 2-chloro-4-nitrophenyl-alpha-D-maltotrioside. Clin. Chem. Lab. Med 37, 1053–1062. https://doi.org/10.1515/CCLM.1999.154.

MacDonald, D., Wetherell, M.A., 2019. Competition stress leads to a blunting of the cortisol awakening response in elite rowers. Front. Psychol. 10, 1684. https://doi.org/10.3389/fpsyg.2019.01684.

Muth´en, L.K., Muth´en, B.O., 2017. Mplus: Statistical Analysis with Latent Variables: User's Guide.

Martin, A.J., Ginns, P., Papworth, B., 2017. Motivation and engagement: Same or different? Does it matter? Learning and Individual Differences 55, 150–162. https://doi.org/10.1016/j.lindif.2017.03.013.

Matsumura, S., Watanabe, K., Saijo, N., Ooishi, Y., Kimura, T., Kashino, M., 2021. Positive Relationship Between Precompetitive Sympathetic Predominance and Competitive Performance in Elite Extreme Sports Athletes. Frontiers in Sports and Active Living 3, 712439. https://doi.org/10.3389/fspor.2021.712439.

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.

Nater, U.M., Rohleder, N., 2009. Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research. Psychoneuroendocrinology 34, 486–496. https://doi.org/10.1016/j.psyneuen.2009.01.014.

Nater, U.M., La Marca, R., Florin, L., Moses, A., Langhans, W., Koller, M.M., Ehlert, U., 2006. Stress-induced changes in human salivary alpha-amylase activity—associations with adrenergic activity. Psychoneuroendocrinology 31, 49–58. https://doi.org/10.1016/j.psyneuen.2005.05.010.

Nater, U.M., Rohleder, N., Schlotz, W., Ehlert, U., Kirschbaum, C., 2007. Determinants of the diurnal course of salivary alpha-amylase. Psychoneuroendocrinology 32, 392–401. https://doi.org/10.1016/j.psyneuen.2007.02.007.

O'Keefe, P.A., Horberg, E.J., Plante, I., 2017. The Multifaceted Role of Interest in Motivation and Engagement. In: O'Keefe, P.A., Harackiewicz, J.M. (Eds.), The Science of Interest. Springer International Publishing, Cham, pp. 49–67. https://doi.org/10.1007/978-3-319-55509-6_3.

Polivy, J., 1998. The effects of behavioral inhibition: integrating internal cues, cognition, behavior, and affect. Psychol. Inq. 9, 181–204. https://doi.org/10.1207/s15327965pli0903_1.

Pruessner, J.C., Kirschbaum, C., Meinlschmid, G., Hellhammer, D.H., 2003. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology 28, 916–931. https://doi.org/10.1016/S0306-4530(02)00108-7.

Richter, M., Friedrich, A., Gendolla, G.H.E., 2008. Task difficulty effects on cardiac activity. Psychophysiology 45, 869–875. https://doi.org/10.1111/j.1469- 8986.2008.00688.x.

Roelofs, K., 2017. Freeze for action: neurobiological mechanisms in animal and human freezing. Philos. Trans. R. Soc. B Biol. Sci. 372, 20160206. https://doi.org/10.1098/ rstb.2016.0206.

Roelofs, K., Elzinga, B.M., Rotteveel, M., 2005. The effects of stress-induced cortisol responses on approach–avoidance behavior. Psychoneuroendocrinology 30, 665–677. https://doi.org/10.1016/j.psyneuen.2005.02.008.

Rohleder, N., Nater, U.M., Wolf, J.M., Ehlert, U., Kirschbaum, C., 2004. Psychosocial stress-induced activation of salivary alpha-amylase an indicator of sympathetic activity? Ann. N. Acad. Sci. 258–263.

Sapolsky, R.M., Romero, L.M., Munck, A.U., 2000. How Do Glucocorticoids Influence Stress Responses? Integrating Permissive, Suppressive. Stimul., Prep. Actions 21, 35. SAS Institute, 2013. SAS 9.4 Statements: Reference.

Scheiter, K., Ackerman, R., Hoogerheide, V., 2020. Looking at Mental Effort Appraisals through a Metacognitive Lens: Are they Biased. Educ. Psychol. Rev. 32, 1003–1027. https://doi.org/10.1007/s10648-020-09555-9.

Sherman, G.D., Mehta, P.H., 2020. Stress, cortisol, and social hierarchy. Curr. Opin. Psychol. 33, 227–232. https://doi.org/10.1016/j.copsyc.2019.09.013.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: Uses in assessing rater reliability. Psychol. Bull. 86, 420–428. https://doi.org/10.1037/0033-2909.86.2.420.

Silvia, P.J., Mironovov´a, Z., McHone, A.N., Sperry, S.H., Harper, K.L., Kwapil, T.R., Eddington, K.M., 2016. Do depressive symptoms "blunt" effort? An analysis of cardiac engagement and withdrawal for an increasingly difficult task. Biol. Psychol. 118, 52–60. https://doi.org/10.1016/j.biopsycho.2016.04.068.

Treadway, M.T., Buckholtz, J.W., Schwartzman, A.N., Lambert, W.E., Zald, D.H., 2009. Worth the 'EEfRT'? The Effort Expenditure for Rewards Task as an Objective Measure of Motivation and Anhedonia. PLoS ONE 4, e6598. https://doi.org/10.1371/journal.pone.0006598.

Vreeburg, S.A., Hoogendijk, W.J.G., van Pelt, J., DeRijk, R.H., Verhagen, J.C.M., van Dyck, R., Smit, J.H., Zitman, F.G., Penninx, B.W.J.H., 2009. Major Depressive Disorder and Hypothalamic-Pituitary-Adrenal Axis Activity: Results From a Large Cohort Study. Arch. Gen. Psychiatry 66, 617–626. https://doi.org/10.1001/archgenpsychiatry.2009.50.

Vrshek-Schallhorn, S., Avery, B.M., Ditcheva, M., Sapuram, V.R., 2018. The cortisol reactivity threshold model: Direction of trait rumination and cortisol reactivity association varies with stressor severity. Psychoneuroendocrinology 92, 113–122. https://doi.org/10.1016/j.psyneuen.2017.11.002.

Vrshek-Schallhorn, S., Velkoff, E.A., Zinbarg, R.E., 2019. Trait Rumination and Response to Negative Evaluative Lab-Induced Stress: Neuroendocrine, Affective, and Cognitive Outcomes. Cogn. Emot. 33, 466–479. https://doi.org/10.1080/02699931.2018.1459486.

Watkins, M.W., 2018. Exploratory Factor Analysis: A Guide to Best Practice. J. Black Psychol. 44, 219–246. https://doi.org/10.1177/0095798418771807.

Watson, D., Anna, L., Tellegen, A., 1988. Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. J. Pers. Soc. Psychol. 8. Way, B.M., Taylor, S.E., 2010. The Serotonin Transporter Promoter Polymorphism Is Associated with Cortisol Response to Psychosocial Stress. Biol. Psychiatry 67, 487–492. https://doi.org/10.1016/j.biopsych.2009.10.021.

Woody, A., Hooker, E.D., Zoccola, P.M., Dickerson, S.S., 2018. Social-evaluative threat, cognitive load, and the cortisol and cardiovascular stress response. Psychoneuroendocrinology 97, 149–155. https://doi.org/10.1016/j.psyneuen.2018.07.009.

Worthington, R.L., Whittaker, T.A., 2006. Scale Development Research: A Content Analysis and Recommendations for Best Practices. Couns. Psychol. 34, 806–838. https://doi.org/10.1177/0011000006288127.

Wr´obel, M., Finogenow, M., Szyma´nska, P., Laurent, J., 2019. Measuring Positive and Negative Affect in a School-Based Sample: A Polish Version of the PANAS-C. J. Psychopathol. Behav. Assess. 41, 598–611. https://doi.org/10.1007/s10862-019-09720-7.