Brain mediators of cardiovascular responses to social threat, Part I: Reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity

Tor D. Wager¹
Christian E. Waugh²
Martin Lindquist³
Doug C. Noll⁴
Barbara L. Fredrickson⁵
Stephan F. Taylor⁶

¹Department of Psychology, Columbia University
²Department of Psychology, Stanford University
³Department of Statistics, Columbia University
⁴Department of Biomedical Engineering, University of Michigan
⁵Department of Psychology, University of North Carolina, Chapel Hill
⁶Department of Psychiatry, University of Michigan

Running Head: CORTICAL MEDIATORS OF SOCIAL THREAT

Please address correspondence to:
Tor D. Wager
Department of Psychology
Columbia University
1190 Amsterdam Ave.
New York, NY 10027
Email: tor@psych.columbia.edu
Telephone: (212) 854-5318
CORTICAL MEDIATORS OF SOCIAL THREAT

Abstract

Social threat is a key component of mental “stress” and allostatic load, and a potent generator of negative emotions and physiological responses in the body. How the human brain processes social context and drives peripheral physiology, however, is relatively poorly understood. Human neuroimaging and animal studies implicate the dorsal medial prefrontal cortex (MPFC), though this heterogeneous region is likely to contain multiple sub-regions with diverse relationships with physiological reactivity and regulation. We used fMRI combined with a novel multi-level path analysis approach to identify brain mediators of the effects of a public speech preparation task (social evaluative threat, SET) on heart rate (HR). This model provides tests of functional pathways linking experimentally manipulated threat, regional fMRI activity, and physiological output, both across time (within person) and across individuals (between persons). It thus integrates time series connectivity and individual difference analyses in the same path model. The results provide evidence for two dissociable, inversely coupled sub-regions of MPFC that independently mediated HR responses. SET induced activity increases in a more dorsal pregenual cingulate region, whose activity was coupled with HR increases. Conversely, SET induced activity decreases in a right ventromedial/medial orbital region, which were coupled with HR increases. Individual differences in coupling strength in each pathway independently predicted individual differences in HR reactivity. These results underscore both the importance and heterogeneity of MPFC in generating physiological responses to threat.
Author note

We would like to thank Niall Bolger for helpful discussion on path analysis, and the authors of SPM software for making it freely available. This paper was made possible with the support of grant funding from NSF 0631637 (T.W.), NIH Grant MH076136 (T.W.), NIH Grant MH59615 (BLF), a pilot grant from the UM fMRI lab to S.T./C.W., positive psychology microgrant to C.W. and dissertation thesis grant to C.W. Author contributions: Design: C.W., B.F. and S.F.T., Data collection: C.W. and D.N., Analysis: T.D.W., M.L., and C.W. Writing: T.D.W., C.W., M.L., S.F.T., and B.F. Matlab code implementing mediation analyses is freely available at
http://www.columbia.edu/cu/psychology/tor/.
Introduction

Human neuroimaging research has focused on the brain’s role in decision-making and action, but much less focus has been given to the brain’s role in regulating peripheral physiology. Understanding brain-body relationships is critical for understanding psychological influences on health, and is central to the fields of neuro-immunology, psychoneuroimmunology, psychosomatic medicine, and aspects of complementary and alternative medicine (Glaser & Kiecolt-Glaser, 2005; R. Lane et al., in press). It is also important for understanding the causes of and effective treatments for diverse medical conditions. These include disorders with an obvious neuro-psychological component, such as depression (Kirsch et al., 2008; Sneed et al., 2008), anxiety (Etkin & Wager, 2007), and pain (Loggia, Schweinhardt, Villemure, & Bushnell, 2008; T.D. Wager, Scott, & Zubieta, 2007), but they also include other areas in which the focus has traditionally been exclusively physiological, including cardiac health (Rozanski et al., 1988), Parkinson’s disease (Benedetti et al., 2004; de la Fuente-Fernandez et al., 2001), asthma (Kemeny et al., 2007), wound healing (Godbout & Glaser, 2006), and resistance to infection (Cohen et al., 2002). Much work has been devoted to understanding psychological effects on health, but very little of it has yet involved detailed analysis of brain circuits and mechanisms in humans.

The brain systems that regulate peripheral physiology have evolved from brainstem-mediated survival-related functions to include cortically-mediated cognitive functions that are likely to be heavily integrated with social and emotional processes in contemporary humans (Bandler & Shipley, 1994; Craig, 2003; Porges, 2003). Thus, understanding brain-body transfer systems may provide clues into the neural organization of social and emotional behavior. A key feature of cognition appears to be the ability to integrate multiple sensory and bodily cues into a central representation of a situation, or a “schema.” A canonical example is the schema “impending threat,” which might be triggered by a combination of visual cues, auditory cues, and conceptual knowledge: Darkness, shadows that look like human forms, the sound of a mechanical click in the silence, and the knowledge that one is walking alone in a dangerous part of the city. As Walter Cannon (Cannon, 1932) described and many others have done since, output from the brain to the peripheral autonomic nervous system and endocrine system prepares us to
respond rapidly and effectively to impending threats. For example, the classic “fight or flight” response involves increases in heart rate, blood flow to the limbs, pupil dilation, slowed digestion, and other changes (Bandler, Keay, Floyd, & Price, 2000; Obrist, 1981). How these cues are integrated, how the schema of threat is represented in the brain, and how the brain produces changes in the body are central questions in the study of psychophysiology. While these are old and important questions that span several fields of human and non-human animal research, they are just now beginning to be addressed using neuroimaging techniques.

There are ample reasons to study physiological threat responses using human neuroimaging. A remarkable feature of humans is that the most powerful elicitors of physiological threat responses in humans are not physical challenges, but social ones. Indeed, social evaluative threat (SET)—the prospect of being judged unfavorably by other individuals in a public setting—produces consistently robust physiological responses in laboratory stress tasks (Dickerson & Kemeny, 2004; Kirschbaum, Pirke, & Hellhammer, 1993). Animal research on the brain modulation of physiology has focused primarily on brainstem centers such as the periaqueductal gray (PAG) (Bandler & Shipley, 1994; Behbehani, 1995), lower brainstem centers (Saper, 2002), and on the amygdala, which is critical for learning the threat value of cues and signaling the PAG to produce anticipatory physiological changes (LeDoux, 2000). These brain centers are critical for physiological and behavioral expression of emotion, and for aspects of emotional learning, but brain processes in these regions are not likely to be sufficient to generate the complex appraisals and schemas that underlie social threat responses (Lazarus, 1966; Tomaka, Blascovich, Kibler, & Ernst, 1997). What is needed are neuroimaging studies that investigate the cortical and higher subcortical regions likely to underlie social threat and performance threat appraisals, and how those regions relate to brainstem areas such as PAG in the generation of physiological responses. A small but growing literature addresses this gap by focusing on social and performance threat states specifically (H. Critchley, 2003; Dedovic et al., 2005; Eisenberger, Taylor, Gable, Hilmert, & Lieberman, 2007; P. J. Gianaros, F. M. Van Der Veen, & J. R. Jennings, 2004; Kern et al., 2008).
Speech preparation as a model for SET

In this study, we assess fMRI activity elicited by speech preparation and its relationship with heart-rate (HR), an integrated measure of sympathetic and parasympathetic innervation of the heart. One of the most potent laboratory stressors involves preparing and giving a speech in front of a panel of critical, expert evaluators. In a review of over 200 studies, Dickerson and Kemeny (Dickerson & Kemeny, 2004) concluded that public speech preparation and performance is arguably the most robust elicitor of hypothalamic-pituitary-adrenal (HPA)-axis responses (e.g., plasma or salivary cortisol) among laboratory tasks, which they link to appraisals of negative social evaluation (“shame”) (Dickerson, Gruenewald, & Kemeny, 2004). Speech preparation is an important aspect of the Trier Social Stress Test (Kirschbaum et al., 1993), a formalized laboratory stressor designed to induce HPA-axis responses.

Preparing and giving a speech also induces robust cardiovascular engagement, including increased blood pressure and heart rate (HR) (Berntson et al., 1994; Cacioppo et al., 1995; Gramer & Saria, 2007; Tugade & Fredrickson, 2004; Uchino, Cacioppo, Malarkey, & Glaser, 1995) that results from both increased sympathetic output and reduced parasympathetic output to the heart (Berntson et al., 1994). Public speaking stressors have produced larger cardiac chronotropic responses than math performance and reaction-time based stressors (Al'Absi et al., 1997; Al'Absi, Bongard, & Lovallo, 2000; Berntson et al., 1994), though HR responses are comparable whether participants are giving a speech before a critical audience or only preparing the speech (Feldman, Cohen, Hamrick, & Lepore, 2004; Gramer & Saria, 2007).

We focused on HR as an outcome measure for several reasons. First, HR increases are robustly elicited by SET, though they vary across individuals (see Figure 1B; (Berntson et al., 1994)). They are substantially more robust than more pure measures of sympathetic and parasympathetic activity collected over short time intervals (Berntson et al., 1994; Cacioppo et al., 1994). Studies of stressor-induced HR reactivity have estimated its internal consistency above alpha = .95 and test-retest reliability around r = .6 after one year (Cacioppo, 1994; Uchino et al., 1995). Second, they can be measured on a roughly second-by-second basis, providing the ability to analyze effective connectivity among key brain regions and HR across time. Third, HR reactivity and cardiovascular
reactivity more generally predict other health-related effects of stressors on the body. Cardiovascular reactivity is heritable (Carroll, Hewitt, Last, Turner, & Sims, 1985) and is correlated with stressor-induced changes in cortisol release (Al'Absi et al., 1997; Lovallo, Pincomb, Brackett, & Wilson, 1990) and immune function (Cacioppo, 1994; Cacioppo et al., 1995; Sgoutas-Emch et al., 1994; Uchino et al., 1995). Finally, HR reactivity and related cardiovascular reactivity measures are risk factors for cardiac dysfunction and mortality (Thayer & Lane, 2007). Short-term social stress has been shown to induce cardiac ischemia (Rozanski et al., 1988) and other clinically relevant measures of cardiac dysfunction, such as left ventricular ejection fraction (Jain, Joska, Lee, Burg, & Lampert, 2001; Jain et al., 1998), which have been shown to predict fatal and non-fatal cardiac events over a 5-year follow-up (Jiang et al., 1996).

**Cortical and subcortical systems linked to threat responses**

The most likely locations for brain generators of autonomic and endocrine responses to SET are in the medial prefrontal cortex (MPFC), which projects reciprocally to a set of interconnected “limbic” cortical regions and subcortical nuclei, including the insula, medial temporal lobes, amgydala, ventral striatum (caudate and putamen), mediodorsal thalamus, hypothalamus, and PAG, as well as other important brainstem nuclei (An, Bandler, Ongür, & Price, 1998; Bandler et al., 2000; Bandler & Shipley, 1994; Barbas, Saha, Rempel-Clower, & Ghashghaei, 2003; Hsu & Price, 2007; Kondo, Saleem, & Price, 2003, 2005; Price, 1999; Saleem, Kondo, & Price, 2008). MPFC has been broadly associated with emotional processes (T. Wager et al., 2008), with dorsomedial and pregenual regions linked to PAG activation, and tasks that engage self-evaluation (Northoff et al., 2006).

A small but growing number of neuroimaging studies have investigated the human brain correlates of autonomic and endocrine responses to stressors. The most consistent findings across studies include dorsal cingulate/MPFC responses linked to stress-induced increases in HR and blood pressure (H. D. Critchley, Corfield, Chandler, Mathias, & Dolan, 2000; H. D. Critchley et al., 2003; H. D. Critchley, Tang, Glaser, Butterworth, & Dolan, 2005; P. Gianaros, F. M. Van Der Veen, & J. R. Jennings, 2004; Gianaros, Jennings, Sheu, Derbyshire, & Matthews, 2007; Gianaros et al., 2008b) and cortisol (Eisenberger et al., 2007). More rostral and ventral areas have been associated
with reduced cortisol reactivity (Eisenberger et al., 2007; Kern et al., 2008), implying a role in successful regulation or protection from stress reactivity. These studies mark an important milestone in the interrelation of human brain activity and physiology, and have confirmed and extended findings from animal models implicating the ventromedial prefrontal cortex (vmPFC), lateral orbitofrontal cortex (OFC), anterior cingulate (ACC), and anterior insula (aINS)—the same regions thought to be most critical for emotional appraisal—in physiological responses to social threat.

One limitation is that nearly all of the studies cited above (and most others) have relied exclusively on between-subject correlations to make inferences about brain-physiology relationships. For example, the large study by Gianaros et al. (2004) mapped brain regions in which individual differences in task-evoked heart period changes correlated with individual differences in task-evoked brain activity. Similarly, Eisenberger et al. (2007) related individual differences in cortisol responses to brain activity responses in a separate social exclusion task. Such correlations do not take full advantage of the capability of fMRI to make many repeated measurements of brain activity over time (typically 200-1500 per individual). Thus, they are severely limited in power by the sample size (to their credit, the examples above have some of the largest sample sizes among imaging studies to date (T. D. Wager, Lindquist, & Kaplan, 2007)). In addition, between-subject correlations are subject to a number of confounds related to individual differences in age, neurovascular coupling, brain morphometry, and other variables. Thus, data on brain-physiological coupling across time would provide an attractive complement to these analyses.

In this study, we extend these results by using a new kind of analysis—multi-level mediation effect parametric mapping—that is specifically designed to link experimental manipulations, brain activity, and physiological output in a single path model. A single-level version of the model was used in (T. D. Wager, Hughes, Davidson, Lindquist, & Ochsner, 2008). The advantage of the multi-level model is that it can incorporate both within-subjects longitudinal effects across time and between-subjects effects of individual differences in the same model. Thus, it can provide inferences about social threat – brain activity – cardiovascular pathways both across time
and across individuals. In addition, it can provide tests of mediation that standard general linear model-based analyses cannot.

We experimentally manipulated SET by asking participants to silently prepare a speech under time pressure (Figure 1). Participants believed that they would have to give their speech, which would be audiotaped during scanning and judged later by fellow students. We monitored HR continuously during fMRI imaging, and our analyses focused on establishing pathways that link the experimental SET manipulation with variations in brain activity and HR.

The overall inference that a region is critical for generating HR responses to SET includes tests at two levels of analysis. The first level of analysis tests associations between SET, brain activity, and HR across time within individuals. At this level, a region involved in generating HR responses to threat should show the following three characteristics. Activity in a brain region should: 1) increase (or decrease) in response to the SET challenge (Path a in Figure 2); 2) predict HR changes over time, controlling for the SET manipulation (Path b); and 3) Mediate the SET-HR covariance. This latter criterion can be evaluated using a mediation test, which formally tests whether the brain region explains a significant proportion of the SET-HR covariance. The second level of analysis concerns HR reactivity. If there are true individual differences in HR reactivity, and if a particular brain region is a mediator of the SET-HR relationship, then the first-level a and b path strengths should be predicted by HR reactivity. That is, for those who show robust HR increases to the SET challenge, brain activity in mediating regions should be more strongly associated with both SET and HR. Inferences about brain regions that link social threat with autonomic activation draw on each of these five hypotheses (three related to dynamic co-variation across time and two related to individual differences.)

Methods

Participants

Thirty healthy, right-handed, native English speakers were recruited at the University of Michigan (mean age 20.3 years, 10 males) and participated in this experiment. Potential participants were initially pre-screened for scoring in the upper or
lower quartile of an emotional resilience measure (ER-89) (Block & Kremen, 1996). However, none of the results presented in this paper were related to this personality trait (p > .5), so the two subgroups were combined in all analyses. Resilience-related results from this sample on a different task are presented elsewhere (Waugh et al., 2008).

Participants were excluded who reported a prior history of neurological or psychiatric illness, current or prior psychoactive medication, claustrophobia, or other standard contraindications for fMRI, and were asked to abstain from tobacco and caffeine use 24 hours prior to scanning. All participants gave written informed consent in accord with the Declaration of Helsinki and as approved by the University of Michigan institutional review board. Two participants were excluded due to excessive head motion (> 3 mm), two were excluded because sufficient anatomical warping quality could not be achieved, and two additional participants did not have complete HR data, leaving a final sample of 24 participants.

**Procedure and fMRI task design**

A schematic description of the 7-min long task is depicted in figure 1A. After an initial anatomical scan, participants were instructed that in a couple of minutes they were to prepare a speech that would be audiotaped in the scanner and then judged by fellow students on persuasiveness, organization and intellectual quality. Participants were given headphones with a microphone attached to add to the illusion that they would be giving a speech in the scanner. They were also told that there was a slight possibility that they would not have to give the speech. After reading these instructions, the participants were told to relax and try to clear their minds of all thoughts and feelings while fixating on a fixation cross for two minutes, during which we acquired baseline physiological and fMRI data. At the end of two minutes, the speech topic “Why am I a good friend” was presented for 15 seconds, and the participants were told they would have two minutes to prepare their speech. After a 2-min preparation period, participants were instructed that they were randomly selected to not give a speech, and asked to relax for the remaining 2.5 minutes of fMRI scanning. We adopted a standardized procedure by presenting computerized instructions during all phases, using E-prime software (PST Inc.).
This is an unusual design because it involves only a single, relatively sustained SET challenge. Allowing participants a sustained 2-min period to develop emotional and physiological responses was a key feature of our design. Because most people cannot easily switch rapidly among emotional states, traditional block or event-related designs dilute emotional experience in favor of repeated occurrences of artificially manipulated epochs. The present task employed a novel analysis method especially suited for capturing brain activity reflecting a fuller, more ecologically-valid emotional experience.

**Data acquisition and preprocessing**

Heart rate was collected continuously with a sampling rate of 100 Hz during entire scanning time, using photoplethysmography on the left index finger. Using customized software from the James Long Company, we first identified and removed artifacts from the data (blind to task condition; outliers were identified with a custom algorithm that estimated improbable inter-beat intervals). Inter-beat intervals were then calculated from the remaining R-waves, and HR was averaged into 2 s bins (the scan repetition time, TR). HR reactivity was calculated as each participant’s mean HR during speech preparation (from the presentation of the speech topic to the presentation of the “relax” cue), compared with pre- and post-stress rest periods.

MRI images were collected on a 3.0T GE whole-body scanner (GE Medical Systems). Structural images were acquired using high-resolution T1 spoiled gradient recall images (SPGR) for anatomical localization and warping to a standard space. Functional blood-oxygen-level-dependent (BOLD) images were acquired with a T2*-sensitive spiral in-out sequence (Glover & Law, 2001) (TR = 2000 ms, TE = 40 ms, flip angle = 90°, 24 slices in ascending sequential sequence, $4.5 \times 3.4375 \times 3.475$ mm voxels). An LCD projector displayed stimuli on a back-projection screen placed in the scanner room.

Functional images were subjected to a standard preprocessing sequence. Slice-timing acquisition correction using sync interpolation was performed using custom software written by Dr. Luis Hernandez, and realignment of the functional images to correct for head movement was performed using the Automated Image Registration tools (Woods, Grafton, Holmes, Cherry, & Mazziotta, 1998). The remaining preprocessing
steps were performed using the Statistical Parametric Mapping analysis package (SPM2, Wellcome Department of Cognitive Neurology, London, UK). SPGR images for each participant were coregistered to the mean functional image. SPGR images were normalized to the anatomical space of the 152-brain template provided by the Montreal Neurological Institute (MNI), and parameters were applied to the functional images. Finally, the normalized functional images were smoothed with an 8-mm full-width at half-maximum Gaussian smoothing kernel. This set of data is referred to as “raw” data in analyses and figures.

Prior to analysis, multiple regression was used to estimate and remove linear effects of a number of known nuisance covariates (see Figure 1C). These included, for each participant, a) 6 estimated head movement parameters from realignment, their mean-zeroed squares, their derivatives, and squared derivatives; b) the whole-brain global signal time series (Vincent et al., 2006) (note that this is not the same as “global scaling” in SPM software and does not suffer from the same problems because it is not a re-scaling of the data (Aguirre, Zarahn, & D'Esposito, 1998)); c) indicator vectors for outlier time points identified based on their multivariate distance from the other images in the sample; and d) linear drift across time. These covariates were removed prior to analysis to ensure that the main analyses were not confounded by these variables. Global outliers (c above) were identified by computing both the mean and the standard deviation of values in each image for each slice. Mahalanobis distances for the matrix of mean values (one per slice) x functional volumes were computed, and images with a value above 3 standard deviations were considered outliers. The same procedure was used for standard deviation values. Typical numbers of outliers ranged between zero and four images per participant.

All analyses described below were conducted on data after removing movement, outlier, and drift-related covariates (see Figure 1C) from each subject’s time series data. Analyses conducted separately on the “raw” preprocessed data showed similar, but less spatially specific, effects (data not shown).

Statistical analysis: Multi-level path modeling

Multi-level path modeling of fMRI data. A typical fMRI model assesses the relationship between experimental manipulations, reflected in comparisons across trial types, and
brain activity. These relationships are typically analyzed using a two-stage “summary statistics” procedure in which within-subjects effects (i.e., differences in fMRI activity among trial types) are estimated for each subject using a first-level analysis, and between-subjects effects are tested using a separate, second-level model, conducted on the regression slopes or contrast values. This two-step procedure is a simplification of a full univariate mixed-effects model that considers both within-subjects (experimental manipulations) and between-subjects (individual differences) variation in the same model (Beckmann, Jenkinson, & Smith, 2003; Raudenbush & Bryk, 2002). Our analyses incorporate a weighted least squares-based mixed-effects model that considers both sources of variation in the same model, and therefore makes fewer assumptions about homogeneity of variances across subjects.

The multi-level path modeling approach we employed extends the univariate mixed-effects model by including an additional outcome variable, heart rate (HR). The path model thus assesses the standard analysis of experimental effects on fMRI activity, as well as several other effects, in the context of a single structural equation model, including relationships between experimental manipulations ($X$), brain activity ($M$), and HR ($Y$); see Figure 2. In addition, because it is a mixed-effects model, it incorporates both “first-level” (within-subjects) and “second-level” (between-subjects) effects (individual differences in brain responses and HR reactivity). Analyses are still performed using data from each brain voxel in a separate analysis, so that the full path model and effects of interest are tested substituting each brain voxel’s data for $M$. This approach, which we have referred to as Mediation Effect Parametric Mapping (T. D. Wager et al., 2008), retains the flexibility of the statistical parametric mapping approach in locating voxels that show particular effects of interest, but extends it to evaluating effects of interest in a simple structural equation model.

The advantages of using the structural model over a standard GLM approach are that: 1) It can provide tests of mediation effects; 2) It estimates several GLM equations in the context of a single path model, making it easy to localize regions that show a pattern of interest across multiple effects; and 3) within-subject measurement error is taken into account when conducting group analyses, providing increased efficiency if data quality is better for some subjects.
In the models we present here, the initial variable \((X)\) in the path model is an experimentally manipulated social evaluative threat (SET) task (which takes on values of 1 during speech preparation and zero otherwise), and the mediating variable \((M)\) is the time series of brain activity in a single voxel. The SET regressor thus captures the contrast [SET – (Pre + Post Baselines)]. We chose this coding because pre- and post-threat baselines were not reliably different physiologically. We note, however, that the linear drift removal (important for fMRI analysis) prevented us from detecting differences in levels of fMRI activity pre- vs. post-threat. One effect of interest was the effect of SET on brain activity—equivalent to the [SpeechPrep – Baseline] contrast in a standard fMRI analysis)—which we refer to as Path a (see Figure 2).

A second effect of interest is the association between brain activity \((M)\) and HR \((Y)\), which we refer to as Path b. As is standard in directed path models, Path b is assessed while controlling for \(X\) (SET), so that any relationships between brain activity and HR cannot be attributed to the experimental manipulation as a third variable (Baron & Kenny, 1986; MacCorquodale & Meehl, 1948). This analysis would typically be undertaken in a separate “effective connectivity” analysis. Finally, a mediation test provides inference on whether the inclusion of brain activity \((M)\) in the model accounts for a substantial amount of the SET \((X)\) to HR \((Y)\) relationship. This is equivalent to testing the product of the path coefficients \(a*b\). We elaborate on this test below. In the present results, for simplicity, we identify regions that show significant \(a\) and \(b\) paths, indicating a relationship with speech preparation and an independent relationship with HR, and then conduct mediation tests on the resulting regions. However, qualitatively identical results were found when voxel-by-voxel maps were made of the overlap of \(a\), \(b\), and \(a*b\) effects.

The system of linear equations underlying the path model are defined as follows. The first level (within-person) equations are of the form \(Y_j = Z_j \beta_j + \varepsilon_j\), for \(j=1, \ldots N\), where \(Z\) is a design matrix for participant \(j\), \(\beta_j\) is a person-specific vector of regression coefficients, and \(\varepsilon_j\) is a residual vector. Each of the \(Z_j\) matrices in the multi-level mediation model contains an intercept column and one or two regressors of interest, as well as regressors for a set of nuisance covariates \(Q_1 \ldots Q_n\). Adopting the notational convention of (Kenny, Korchmaros, & Bolger, 2003), the equations can be written:
Here the subscript $i$ indexes observation (time point) and $j$ indexes person. The parameters $d_{0,2}$ represent intercept values, and $q_1 \ldots q_n$ represent nuisance regression parameters, neither of which are of further interest in our analysis. The parameters $a$ and $b$ are linear regression slopes, as described above. In addition, $c$ represents the total relationship between $X$ and $Y$ (controlling for nuisance variables), and $c'$ represents the direct relationship controlling for $M$. As we are interested in mediation effects, and complete mediation by any single brain region is unlikely in the present analyses (i.e., a small and non-significant $c'$), $c$ and $c'$ are not discussed in detail in this report. Finally, $r$, $e$, and $f$ represent error terms.

The second-level (between-person) equations are of the general form $\beta_j = \gamma + \nu(HRR) + u_j$, indicating that an individual regression slope ($\beta_j$) is modeled as the sum of a population regression slope ($\gamma$), an effect of a second-level moderating individual differences variable ($\nu$), in this case HR reactivity, and a person-level error term ($u$) for that effect. The group second-level design matrix $Z_g$ contains an intercept column and the HR reactivity covariate, making the second-level structural equations as follows:

$$a_j = Z_g \begin{bmatrix} \gamma_a \\ \nu_a \end{bmatrix} + u_{aj}; \quad b_j = Z_g \begin{bmatrix} \gamma_b \\ \nu_b \end{bmatrix} + u_{bj}; \quad c_j = Z_g \begin{bmatrix} \gamma_c \\ \nu_c \end{bmatrix} + u_{cj}; \quad c_j' = Z_g \begin{bmatrix} \gamma_{c'} \\ \nu_{c'} \end{bmatrix} + u_{c'j}$$

Equations for the intercept terms are not shown, but follow the same form. Thus, person-level intercept and slope estimates are treated as random effects. The within-person error terms ($r_j$, $e_j$, and $f_j$) and the between-person error terms ($u_{0j}$, $u_{1j}$, $u_{2j}$, $u_{aj}$, $u_{bj}$, $u_{cf}$ and $u_{c'j}$) are each assumed to be normally distributed with a mean of zero and a unique, but constant, variance for each effect.

**Model and variance component estimation.** Full mixed-effects models typically iterate between estimation of model parameters and the within-person and between-subject variance components using algorithms such as expectation-
maximization (Dempster, Laird, & Rubin, 1977) or iterative generalized least squares or Fisher scoring (e.g., Goldstein, 1986). Variance component estimates are obtained from the residuals using restricted maximum likelihood (ReML) (Goldstein, 1989; Johnson & Thompson, 1995). However, this procedure can be extremely computationally intensive, and much of the computation time is spent in iterative estimation of how much of the total error variance is attributed to within-subject \((r, e, \text{and } f)\) and between-subject error terms \((u \text{'s})\). Here, we approximate the iterative solution using a 3-step re-weighting. This technique allows us to estimate the multi-level path model over 200,000 brain voxels in a reasonable amount of time (~6 hours on a 2008 Intel Mac 8-core workstation).

Specifically, the following procedures are used separately for each regression equation in the path model (e.g., Eqs. 1, 2, and 3), which are described in more detail below: a) Ordinary least squares estimation of regression parameters and variances; b) Estimation of precision values based on the inverse of naïve total (between-subjects plus within-subjects) variance estimates; c) Precision-weighted updating of regression parameters (re-weighting Step 1); d) Empirical Bayes updating of the between-subjects error variance and precision estimates (Step 2); and e) Estimation of final regression parameters using weighted least squares (Step 3). All of these procedures are implemented in the Multilevel Mediation/Moderation Toolbox (M3) v.0.9 (T.D.W. and M.L), available from the authors.

Initially, first-level slope and intercept parameters and their variances are calculated for each participant using standard ordinary least squares multiple regression. If \(Z_j\) is a design matrix for a single regression equation (i.e., for the \(a, b/c'\), or \(c\) effect) with \(t\) time points (observations) \(x\) \(k\) regressors for subject \(j\), and \(Y_j\) is the \(t \times 1\) outcome data vector, then the \(k \times 1\) vector of regression parameter estimates can be calculated as:

\[
\hat{\beta}_j = (Z_j^T Z_j)^{-1} Z_j^T Y_j
\]

with residual variance

\[
\hat{\sigma}^2_j = (Y_j - Z_j \hat{\beta})^T (Y_j - Z_j \hat{\beta}) / (t - k)
\]

and parameter variance/covariance matrix

\[
V_j = \text{cov}(\hat{\beta}_j) = (Z_j^T Z_j)^{-1} \hat{\sigma}^2_j
\]
For the mediation \((a*b)\) effect, however, we replace Eq. 10 with the formula from Kenny, Korchmaros, and Bolger (2003):

\[
V_j = \text{Var}(ab_j) = b_j^2 \text{var}(a_j) + a_j^2 \text{var}(b_j) + \text{var}(a)^2 \text{var}(b)^2
\]  
(11)

We note that an autoregressive level-1 error structure was not specified in the model because although allowing for autocorrelation can produce maximally efficient inferential statistics, autocorrelation parameter estimates based on noisy level-1 data are not necessarily more efficient (Friston et al., 2000), and parameter estimates are unbiased in either case.

Let \(\hat{B}\) be the \(N \times k\) matrix of parameter estimates across all subjects. An initial estimate of the between-subjects covariance component \((U)\) is the \(k \times k\) matrix:

\[
\hat{U} = \text{cov}(\hat{B}) = (\hat{B}^\top \hat{B})/(N-1)
\]  
(12)

where \(N\) is the number of participants. The precision matrix of each individual participant’s estimates is thus the inverse of the sum of between- and within-subjects error estimates:

\[
\hat{P}_j = (\hat{U} + \hat{V}_j)^{-1}
\]  
(13)

These initial estimates are used to provide weighted least squares estimates of the population regression slope and intercept parameters (the \(k\)-length vector \(\hat{\gamma}\)) (Raudenbush & Bryk, 2002, eq. 3.31):

\[
\hat{\gamma} = \left(\sum_j \hat{P}_j\right)^{-1} \sum_j \hat{P}_j \beta_j
\]  
(14)

The regression slope parameters for each participant are re-estimated using an Empirical Bayes average of the original estimates and \(\hat{\gamma}\) (Raudenbush & Bryk, 2002, eq. 3.56):

\[
\hat{\beta}_j^* = \hat{U} \hat{P}_j \beta_j + (I - \hat{U} \hat{P}_j) \hat{\gamma}
\]  
(15)

The original estimate of \(U\) will usually be an over-estimate, as no attempt is made to factor out within-subjects measurement error as a source of variance. Rather than perform a series of computationally demanding iterations, we use the Empirical Bayes estimates to provide a revised estimate of \(U\) that is shrunk towards zero: \(\hat{U}^* = \text{cov}(\hat{\beta}^*)\). Finally, subject weights are calculated based on the total variance of each parameter estimate for each subject and normalized to sum to 1:
This new estimate of the population parameters is used to update the between-subjects covariance estimate.

These weights are used in the final second-level weighted least squares analysis for each effect of interest ($a$, $b$, $a*b$). For each effect, if $W$ is a diagonal matrix containing the weight values $w_j$ for each participant, then:

$$ w_j = \left( \sum \hat{P}_j^* \right)^{-1} \hat{P}_j^* $$

In this study, the group design matrix $Z_g$ consisted of an intercept and mean-centered HR reactivity predictor. The intercept term in the second-level model is the basis for a test of significance for whether the average $\beta_j$ differs from zero, and thus tests the reliability of Paths $a$, $b$, and $a*b$ across participants. The HR reactivity term tests whether path slopes are predicted by HR reactivity. For Path $a$, a positive finding for HR reactivity implies that HR reactors show a stronger brain response to SET. For Path $b$, it implies that HR reactors show a stronger brain-heart relationship across time. And for Path $a*b$, it implies that the functional pathway from SET to brain to HR is more engaged in HR reactors.

Significance testing and inference for each effect was performed as follows. The residual-inducing matrix $R$ aids in the estimation of variances and degrees of freedom:

$$ R = W^{1/2} (I - Z_g^T W Z_g)^{-1} Z_g^T W $$

The error degrees of freedom are estimated based on the weighting using the Satterthwaite correction (Satterthwaite, 1946):

$$ df = \text{trace}(RW^{-1})^2 / \text{trace}(RW^{-1}RW^{-1}) $$

and the parameter variance/covariance matrix is:

$$ \text{cov}(\hat{\gamma}) = \frac{\left(R\hat{\beta}_j\right)^T}{df} \left(R\hat{\beta}_j\right) \left(Z_g^T Z_g\right)^{-1} $$

Standard $t$-values are obtained for purposes of inference using the equation:

$$ t_{r,k} = \frac{\hat{\gamma}_k}{\text{cov}(\hat{\gamma})^{-1}} $$
Uncorrected P-values are obtained from this t-test (with, in this case, 21 degrees of freedom). In order to correct for multiple comparisons, we calculated the experiment-wise false discovery rate (FDR) (Genovese, Lazar, & Nichols, 2002) at $q < .05$ across all effects of interest, including the first- and second-level $a$, $b$, and $a*b$ coefficients (equivalent to $P < .00037$ uncorrected).

**Analysis strategy.** In this paper, we used a whole-brain, voxel-wise search to identify regions that showed overlapping Path $a$ (SET responses) and Path $b$ (HR connectivity) effects at the first level (Figure 2). We also report results from voxel-wise search for second-level, HR reactivity effects. Thus, inferences localizing regions related to both SET and HR ([Path $a$, Path $b$] overlap regions, Figure 3) were drawn from a multi-voxel brain mapping approach, with correction for multiple comparisons.

We conducted additional analyses on first-level [Path $a$, Path $b$] overlap regions and amygdala regions of interest (ROIs) defined *a priori* to characterize the nature and pattern of effects in these regions (Figure 3). We identified four sub-regions of the amygdala defined by Amunts et al. (Amunts et al., 2005) and labeled in the SPM anatomy toolbox v1.5 (Eickhoff et al., 2005), including the left and right corticomedial group (superior amgydala) reported most frequently in imaging studies of emotion (T. Wager et al., 2008) and the larger left and right basolateral complex.

Within these regions, we conducted three kinds of analyses. For these analyses, fMRI data were averaged over contiguous voxels in each region within each participant, yielding a single brain time-series per region per participant. The first analysis was intended to test whether brain-HR connectivity could be explained by some general characteristic of task demand over time that was not captured in our linear SpeechPrep regressor, such as a high demand on internal vocalization early in the preparation period (Figure 4). We reasoned that if brain activity correlated more strongly with individual HR time series than with the group, it would be less likely that task demands common to the group could have created the observed results. Details are presented below.

Secondly, we conducted the full multi-level path model on each region. For the [Path $a$, Path $b$] overlap regions, this analysis simply duplicated the Path $a$ and Path $b$ effects used to select voxels (thus, the results are merely descriptive), but the model provided additional new information on the mediation ($a*b$) effect (see below) and on the...
second-level relationships with HR reactivity. We tested separate path models that included only one brain region as a mediator (M) variable (Figure 5). We also tested a single path model that included multiple key [Path a, Path b] overlap regions in the same model (Figure 6). In this model, Eq. 3 is altered so that additional mediating regions are controlled for when assessing each of the Path b (brain-HR) effects. Thus,

\[ Y_j = d_{2j} + c_j X_j + b_{1k} M_{1ij} + b_{2k} M_{2ij} + \ldots + b_{kj} M_{kj} + q_{ij} Q_{ij} + \ldots + q_{nij} Q_{nij} + f_j \]  

(22)

The results from both types of model were qualitatively identical, in that the significance of the path coefficients was not affected by whether other mediators were included in the model. Thus, we focus on the results of the combined path model, which demonstrates independent mediation by multiple brain regions.

Third, we conducted cross-correlation analyses of the relative timing of the brain-heart relationship, to determine if the estimated lag of brain-HR connectivity was earlier for ventral MPFC than dorsal MPFC. Parasympathetic effects on HR are immediately apparent but sympathetic effects are delayed by at least 2 sec and probably longer (Berntson, Cacioppo, & Quigley, 1993). Therefore, earlier brain-HR latencies in VMPFC would be consistent with the notion that ventral MPFC is more closely associated with parasympathetic control (P. Gianaros et al., 2004). Cross-correlation analyses identified a lag \( \delta \) and its standard error across participants that maximized the brain-HR correlation. Details of this method are described below.

**Interpretation of the multilevel mediation (\(a*b\)) effect**

The mediation test provides additional information beyond testing whether the conjunction (intersection) of both Path a and Path b effects are significant. In a single-level path model, the mediation test is significant in a subset of models that show both a and b effects. In a single- or multi-level path model, both a and b paths can be significant without a significant \(a*b\) mediation effect under two conditions: First, the results for one path or the other may be too weak, even if the other is highly significant. Secondly, in a multi-level model, mediation implies that the a and b paths are functionally linked. The strength of the total mediation effect \(a*b\) equals (see Kenny, Korchmaros, & Bolger 2003, Eq. 9):

\[ c - c' = \text{mean}(a*b) = \text{mean}(a) \times \text{mean}(b) + \text{Cov}(a, b) \]  

(22)
Hence, the $a*b$-effect that we report in this paper can be driven by two different sources, the product of the means of $a$ and $b$ and the covariance between $a$ and $b$. The first source comes from first-level mediation effects that are replicable in the group, and the second source comes from individual differences in coherent pathway strength. In the present study, we report only regions that show significant results on all effects ($a$, $b$, and $a*b$), implying first-level mediation effects generalizable to the population, though coherent individual differences in $a*b$ pathway strength may also contribute.

**Correlations with own vs. average HR**

For each region analyzed and for each participant $j$, we calculated the correlation between the participant’s brain time series and that participant’s HR time series. We also calculated the correlation between the brain time series and the group-average HR time series. We converted Pearson’s $r$ values to Fisher’s $Z$ scores and performed a paired t-test on the $Z$ values for the participants’ own vs. group-average HR correlations.

**Cross-correlation analysis**

Brain-HR timeseries were cross-correlated by shifting the brain time series ($Y_j$) for each participant by lag $\delta$ relative to the HR time series ($H_j$) for that participant using a fast frequency-domain-based shift (M.A.L.) at a resolution of 0.4 sec intervals. All intervals between $-8$ sec and $+8$sec were evaluated, using the correlation p-value as an outcome. After this evaluation, the $\hat{\delta}$ yielding the maximal $P$-value was used as a starting value in a nonlinear fitting method estimating the optimal $\delta$, constrained to a window of +/- 2 sec around the initial $\delta$ estimate. The objective function for this estimation was $P_j = f(Y_j, H_j, \delta_j)$, where $f(x)$ was an inferential test of the cross-correlation at lag $\delta$. This procedure was performed for each subject, yielding estimates of $\delta_j$. To further stabilize group estimates, a second-pass analysis was performed. We used Empirical Bayes priors with a normal distribution of mean 0 and standard deviation

$$\left( \sum_j (\hat{\delta}_j - \bar{\delta}_j)^2 / (J-1) \right)^{1/2},$$

and re-computed the nonlinear estimation using the maximum a posteriori (MAP) estimate of $\hat{\delta}_j$ based on both the estimated P-value and the prior distribution. The Bayesian analysis thus shrinks estimates ($\hat{\delta}_j$) toward zero, but reduces the variance associated with the estimation.
Final statistics on whether the $\delta$s different significantly from zero were performed using a sign permutation test (M. A. Lindquist, Loh, Atlas, & Wager, in press), a semiparametric test that does not assume a normal error distribution. Negative $\delta$ values indicated that the brain time series preceded the HR time series, and positive $\delta$ values indicated that the brain time series lagged behind the HR time series. However, we note that no model was used to account for the hemodynamic lag in fMRI signal; thus, $\hat{\delta}$ values near zero suggest that the actual neural activity driving the fMRI signal happened on the order of 5-6 sec before the observed change in HR. We do not attempt to quantify this more precisely because the hemodynamic response shape and latency varies across brain regions and tasks.

Results

Physiological effects of SET
HR changes over time were a primary outcome measure of the SET challenge. Compared with pre-threat baseline and post-threat recovery, speech preparation induced reliable increases in HR (9.54 beats per minute, BPM, $t = 4.88$, $p < .0001$), as shown in Figure 1B. This effect was significant when preparation was compared separately with each of the first (8.16 BPM, $t = 4.18$, $p = .0004$) and second baseline periods (10.92 BPM, $t = 5.27$, $p < .0001$). HR was lower during the post-preparation recovery period than during the first baseline (2.76 BPM, $t = 2.91$, $p = .008$). The same relationships were found after regressing out all MRI-related covariates within each subject ($p \leq .0003$ for all tests), except that the first and second baseline periods were no longer significantly different (0.07 BPM, $t = 0.91$, $p > .10$). There were 22 degrees of freedom for each planned test. HR reactivity was uncorrelated with ego resilience ($r = .01$, $P > .5$) or other measures collected in the sample reported in Waugh et al. (2008), including sex, ethnicity, income, age, and several other trait personality measures (all $P > .09$).

In the brain analyses, the effect of Speech Preparation Status (Speech Prep) was coded in the model as a regressor with values of 1 during speech preparation, and values of zero during pre-threat baseline and post-threat recovery. This regressor thus captures the contrast [SET – (Pre + Post Baselines)]. We note, however, that the linear drift removal (important for fMRI analysis) prevented us from detecting differences in levels
of fMRI activity pre- vs. post-threat. The threshold for all effects was $q < .05$ FDR-corrected for multiple comparisons ($P < .00037$). Coordinates and peak Z-values are listed in Table 1.

**Brain responses to SET (Path a)**

Figure 2 shows a path diagram that links Speech Prep (the experimental manipulation of SET), brain activity, and HR variation. The first within-subjects effect we tested localized brain regions that responded to the SET challenge, compared with the average of first and second baseline periods (Path $a$ in Figure 2). This contrast was expected to show significant responses in a broad set of regions, including those critical for generating autonomic responses, visual responses to task instructions, and others.

Increases in activity to Speech Prep were found most prominently in the MPFC, and in particular the pre-genu anterior cingulate cortex (pgACC), extending into the right caudate, as well as several temporal and occipital cortical regions (yellow/red in Figure 2A). All significant regions are reported in Table 1. Decreases in activity to the SET were found prominently in the OFC, as well as the right putamen, ventral anterior insula, temporal cortex, and posterior cingulate cortex.

The Path $a$ effect at the second level tested the relationship between brain responses to speech preparation and HR reactivity—that is, the correlation between individual brain activity scores and HR reactivity scores for the [SpeechPrep – Baseline] contrast. The results are shown in Figure 2C and Table 1. Positive effects, indicating positive correlations with HR reactivity, were found in a subset of regions showing first-level effects, including pgACC, extending into the right caudate nucleus, and right ventral temporal cortex (shown in warm colors). Negative correlations between SET and HR reactivity were found in bilateral OFC and right putamen (shown in cool colors).

**Brain correlates of HR (Path b)**

The second test for identifying brain regions that mediate the response of threat on HR is to identify regions that predict HR changes, both across time (first-level) and at the individual differences level (second-level). The brain-HR relationship was tested controlling for the Speech Prep indicator—a standard technique in path modeling to establish a mediator-outcome (here, brain-HR) relationship independent of task.
demand—and is depicted as Path $b$ in Figure 2. At the first level, brain time series were correlated positively with HR variations across time most prominently in the pgACC, extending into the rdACC, VMPFC, and right caudate head, as well as in the left DLPFC (middle frontal gyrus) and right caudate body. Negative correlations with HR were found in the right medial and lateral OFC, right inferior frontal gyrus, right putamen and ventral insula, and bilateral temporal pole and parahippocampal cortex. Coordinates and statistics for all Path $b$ effects are reported in Table 2, and group-averaged time series data for key regions are shown in Figure 2.

At the second level, the slope of brain-heart connectivity (Path $b$) was positively moderated by HR reactivity in several regions, including the pgACC, rdACC, and left caudate (Table 2). That is, in these regions, the magnitude of an individual’s HR reactivity to Speech Prep was correlated with the strength of the brain-HR coupling (controlling for SET). All of the regions listed above showed reliable first-level effects as well, indicating reliable brain-HR correlations in the group as a whole. In addition, another set of regions showed second-level moderation but did not show reliable brain-HR (Path $b$) correlation in the group. These regions may thus be associated with positive brain-HR connectivity in high HR responders, but negative brain-HR connectivity in low HR responders. These regions included the pre-SMA, superior frontal cortex, and additional areas of the right caudate.

Negative second-level correlations between brain-heart connectivity and HR reactivity included several areas showing negative brain-heart correlations at the first level: right OFC, right putamen, and left temporal pole (Table 2). Thus, these regions showed a negative time series correlation (Path $b$) that was stronger (more negative) in high HR responders. Additional negative moderation was found in other regions that did not show significant first-level effects, including left OFC, VMPFC, several temporal cortical regions, and brainstem regions including the midbrain and hypothalamus. Because the second-level moderation analysis tests the relationship between brain-heart regression slopes and HR reactivity, these regions showed negative brain-heart associations in the high HR responders, and positive brain-heart associations in the low HR responders.
**ROI analysis of the amygdala**

Notably, the pattern of results described above did not include the amygdala, which was recently implicated in a performance threat by Gianaros et al. (Gianaros et al., 2008b) and has been a focus of attention in studies of conditioned fear and other negative affective processes (Etkin & Wager, 2007; Phelps, Delgado, Nearing, & LeDoux, 2004b; T. Wager et al., 2008; T. D. Wager et al., 2008). We performed ROI analyses on anatomically defined amygdala subregions (see Methods) including the left and right corticomedial group (superior amygdala) and basolateral complex. No positive associations with SET or HR (Path a or b) at either the first or second levels were found in any regions. All subregions showed significant or near-significant de-activations during Speech Preparation (negative Path a effects; see Table 3). Only in the right basolateral amygdala was brain activity predictive of HR, as indicated by negative Path b effects at first- and second-levels (Table 3). A trend towards significant mediation (a*b) was found in this area. Significant mediation was found in the right corticomedial group, but in the absence of significant average a and b effects, this finding indicates the presence of a reliable pathway that differs across individuals in whether increases or decreases lead to HR increases. The group-average time course plot from the right basolateral amygdala is shown in Figure 2. Overall, the results suggest that amygdala increases are not a feature of speech task-related SET, differentiating it from the performance threat results of Gianaros et al. (2008) and other paradigms that elicit negative emotional experience, such as conditioned fear.

**Overlap of SET and HR effects (Path a and Path b conjunction)**

The key regions that link social threat with HR changes should show both SET responses (Path a) and HR correlations controlling for SET (Path b; see Methods). Figure 3 shows the regions with significant results in both tests (at q < .05 FDR in each), along with their group-average time series (using methods described in (Martin A. Lindquist, Waugh, & Wager, 2007); see the figure legend for details). Statistical details are listed in Table 4. In Figure 3, regions showing positive first-level Path a and b effects in red, and negative first-level Path a and b effects in blue. The pgACC and right caudate head were the only regions that showed positive effects in both tests. The right OFC and putamen
showed negative effects in both tests, indicating that they are de-activated by SET and that the degree of de-activation predicts HR increases across time.

Thus, the pattern of results presented thus far suggests that there are bi-valent effects of SET that are linked with HR increases, including activation in the pgACC and de-activation in the OFC and striatum. Notably, each of these regions showed second-level individual differences effects consistent with the first-level time series analyses. Strong HR “reactors” showed larger (more positive) SET-pgACC and pgACC-HR effects, and also larger (more negative) SET-OFC, OFC-HR, SET-putamen, and putamen-HR effects (see Tables 1 and 2).

**Does the brain track individual profiles of HR changes across time? Detailed analysis of pgACC, OFC, and putamen**

We found reliable relationships between brain and HR in several key regions (pgACC, mOFC, and right putamen), both in time series correlations and in correlations with HR reactivity. However, there are a number of potentially confounding processes, such as internal sub-vocalization or cognitive planning activity, that are potentially correlated with HR and not captured completely by the box-car Speech Prep indicator variable. Thus, to more firmly establish the relationship between brain activity and HR, it is important to test how tightly coupled brain time series are with individual patterns of HR.

We first visualized the data in the key regions to assess the strength of the relationship between brain time series and HR changes across time. Detailed plots of the data for the pgACC are shown in Figure 4. Figures 4A and 4B show superimposed plots of the group-average brain and HR time series for the raw data and with covariates removed, respectively. Other regions showed similar data quality, but are not shown for space reasons. In addition, the relationship was apparent for individual participants, as shown in Figure 3C.

Because individual HR responses varied significantly across individuals, another prediction about mediators of the brain-HR relationship is that the brain activity for an individual should track that individual’s specific profile of HR changes across time. To test this hypothesis, we tested whether each participant’s brain time series was more
strongly correlated with the individual’s own HR time series than with the other participants’ time series (correlations were computed for each other participant and averaged). The group-average HR can be considered one kind of surrogate measure related to task demands that vary across time but are common to the group, such as the tendency to engage in cognitive planning early in the speech preparation session but not later. Figure 3D shows a bar graph of the results for the pgACC, showing reliably stronger correlations for the own time series ($Z_{\text{diff}} = 0.18$, $t(22) = 3.11$, $p = 0.005$).

Correlations in the OFC and putamen were also stronger (more negative) for each individual’s own HR profile than others’ (mOFC: $Z_{\text{diff}} = -0.15$, $t(22.0) = -2.90$, $p = 0.008$; putamen: $Z_{\text{diff}} = -0.10$, $t(22.0) = -2.92$, $p = 0.008$) This analysis helps to argue against the notion that the observed brain-HR correlations are a side effect of brain responses to some general type of task demand that is consistent across subjects but not captured by the linear SpeechPrep regressor. But more importantly, perhaps, it simply shows that brain-HR relationships hold even though the HR response profile across time varies across individuals. This implies that individual variability over time is reliable (not just measurement noise).

**Multilevel mediation analyses of key regions**

The full multi-level path model was tested on regions that showed both first-level SET (Path a) and HR (Path b) effects. This analysis provides only descriptive reports of statistics for the first-level $a$ and $b$ effects used to select voxels (since inferences on these regions was already provided by the voxel-wise search), but it provides new inferences on the significance of the formal mediation ($a*b$) effect and the second-level relationships with HR reactivity. The test of mediation amounts to a test of whether controlling for each brain mediator explains a significant amount of the co-variance between the SpeechPrep predictor and the HR time series (see Methods).

We tested both path models with only a single brain mediator (the pgACC, mOFC, or putamen) and a path model with all three regions entered as mediators in the same path model. The significance of the individual regions’ path coefficients was qualitatively the same in both cases. For illustrative purposes, we show plots and statistics for the path model with the pgACC as the only mediator in Figure 5. The left panels (A) shows plots for the first- and second- level results for Path $a$ ($P < .001$ at both
levels), and the right panels (B) shows plots for Path $b$ ($P < .001$ at both levels). The line plots in Figure 5 show first-level regression slope estimates for each participant as blue lines, and the group average with its standard error in gray. The second-level moderation scatterplots show the relationship between path amplitude and HR reactivity. Notably, the mediation effect was significant ($a*b = 0.29, t(21) = 5.22, P < .001$), and second-level analyses showed that HR reactivity positively predicted Path $a$, Path $b$, and Path $a*b$ slopes. That is, the functional pathway from SET to brain to HR was strongest in those with high HR reactivity.

The path model with all three regions entered as mediators is shown in Figure 6. This analysis was designed to distinguish between two alternative hypotheses. First, if the pgACC, OFC, and putamen are independent mediators of threat effects on the heart, then they should be significant mediators even when other regions are included (and their activity controlled for) in the model. Alternatively, activity in these regions could be part of a single bi-valent pattern or distributed “mode”; that is, each region may carry redundant information. If the dorsal and ventral MPFC are relatively specialized for sympathetic and parasympathetic control, respectively, as previous analyses have suggested (P. J. Gianaros et al., 2004; R. D. Lane, Reiman, E.M., Ahern, G. L., & Thayer, J.F., 2001), then we might expect the first case.

The results showed that each of the brain regions was an independent mediator of SET-HR covariance. As listed in Table 5, first-level $a$ and $b$ path coefficients remained significant for each of the three regions ($P < .05$ in all cases), and the first-level mediation effect was significant for each region as well ($P < .005$ in each). Thus, each region independently mediated some of the relationship between experimentally manipulated SET and HR profiles across time. At the second level, HR reactivity moderated the strength of all first-level effects ($P < .05$ in all cases), with the sole exception of a non-significant moderation of the putamen-HR partial path coefficient. Thus, the degree to which a participant was an overall autonomic “responder” moderated each of these functional SET-brain-HR pathways. Partial regression plots are shown in Figure 6 for the pgACC and mOFC.

In Figure 6, significant effects are shown as black arrows, and non-significant effects shown as light gray arrows. The intrinsic “effective connectivity” among these
three regions was assessed by regressing each brain region on the others using the multi-level path modeling framework (see Methods), controlling for indirect effects of the SpeechPrep predictor and fMRI time series from other regions. These analyses showed that the pgACC was functionally coupled with both mOFC ($t = -6.53, p < .0001$) and putamen ($t = -6.26, p < .0001$), but mOFC and putamen were not significantly coupled ($t = 1.50, p = .15$). This suggests that there are functional relationships between the regions that show activation and deactivation in response to SET, but that these relationships are not sufficiently strong that the information about HR contained in the regions is redundant.

**Analyses of timing: Relative latency of brain-heart connectivity**

We used cross-correlation analyses to identify whether there was evidence that the fMRI time series in each region was lagged relative to the HR time series (see Methods). One reason for doing this is simply to determine if our predictions of HR would be systematically improved by taking into account the hemodynamic lag in the fMRI response (no hemodynamic model was assumed in the path analyses). Another reason is to provide a rough, indirect test of the relative contributions of parasympathetic and sympathetic control in each region. Parasympathetic effects on HR are immediately apparent (e.g., vagal stimulation can delay the occurrence of the next heart beat) but sympathetic effects have an estimated onset delay around 2 sec and slower decay constants on HR (Berntson et al., 1993). We would expect earlier brain activity to predict HR in regions associated with parasympathetic control.

Previous studies have associated mOFC and VMPFC with measures of parasympathetic activity (high-frequency heart rate variability) (P. Gianaros et al., 2004; R. D. Lane, Reiman, E.M., Ahern, G. L., & Thayer, J.F., 2001). Consistent with this finding, we found that the cross-correlation between mOFC and HR was marginally stronger when the brain time series preceded the HR time series, with a lag of (-310 ms, SE = 200 ms, $p = .13$). The mOFC lag with HR was significantly earlier than the pgACC lag with HR (-323 ms, SE = 136 ms, $p = .02$). The lag was not significantly different from zero in other regions. The estimated lags for mOFC and other regions are shown in Figure 8. We note that because of the Bayesian lag estimation procedure, the estimated
lags are biased towards zero, so their magnitudes are not directly interpretable in terms of biological time constants.

Overall, the results support the hypothesis of bi-directional control of HR by dorsal and ventral MPFC sub-regions, with HR increases associated with pgACC/rdACC increases and mOFC decreases. Amygala activity decreased with SET and decreases were predictive of HR increases. A dissociation was also found in the striatum, with increases in the caudate head and decreases in right putamen associated with HR increases.

Discussion

Social status and perceived social and intellectual competence are extremely important factors in modern human life (Fiske, Cuddy, & Glick, 2007). Threats to social status and negative evaluation are among the most potent laboratory and real-life stressors in contemporary society. Social evaluative threat (SET) triggers reliable patterns of autonomic reactivity, which can be reliably indexed by heart rate, cortisol, and immune changes. Because SET often arises from a complex analysis of inter-personal relationships, rather than the presence of any particular simple sensory cue, it is likely to be generated by high-level appraisal processes involving the frontal cortex. Thus, SET is a good model for studying the cortical and subcortical brain networks associated with peripheral physiological changes. We used silent speech preparation, a component of a common laboratory stressor that produces robust HR responses (Berntson et al., 1994; Kirschbaum et al., 1993) and other clinically relevant cardiac events (Rozanski et al., 1988), as a SET manipulation during scanning.

In this study, we have found evidence for multiple independent mediators of social threat effects on HR. Specifically, we found that a particular pattern of cortical activity is tightly coupled with HR reactivity to speech preparation. This pattern includes reciprocal activation changes in two sub-regions of the MPFC—pgACC and vmPFC/mOFC—and the right putamen, a portion of the striatum. SET was associated with increases in pgACC (Path a), and those increases predicted HR variations across time (Path b). SET was associated with decreases in vmPFC/mOFC and striatum (negative Path a), and the larger the brain decreases, the higher the HR (negative Path b). These effects at the first level of analysis (time series) suggested that changes in brain
activity predict the changes in HR across time, even controlling for the task period (Baseline, Speech Prep, and Recovery).

In addition, individual differences in SET-brain and brain-HR coupling predicted the magnitude of HR reactivity. This is evidenced by results from the second level in the multi-level path model. Stronger HR “reactors” expressed the pattern of increases in pgACC and decreases in vmPFC and striatum more strongly than non-responders. These effects are captured in the second level moderation of Path a effects by HR reactivity. Stronger HR “reactors” also showed stronger coupling between pgACC increases and HR, and between vmPFC decreases and HR. These effects are captured in the second level moderation of Path b effects by HR reactivity. Overall, the pattern is consistent with the notion that some individuals did not show strong reactions, and thus showed little task-driven variability in both brain activity and HR. Combined, the within-subjects and between-subjects results provide evidence for a dual-process model of HR control in different medial prefrontal sub-regions.

The pgACC is a distinct subdivision of the anterior cingulate from the more ventral subgenual region and the more dorsal anterior mid-cingulate cortex according to Vogt (Palomero-Gallagher, Vogt, Schleicher, Mayberg, & Zilles, 2008; Vogt, 2005). The pgACC and rdACC are together associated with diverse emotional processes (T. Wager et al., 2008), such as in the context-driven modulation of emotion and pain (Petrovic et al., 2005; Porro, 2003; T.D. Wager et al., 2007)(Faymonville, submitted). In a companion paper (T. D. Wager et al., submitted), we have replicated this basic finding in a separate sample, and have extended these results with several additional analyses, including analyses of brainstem mediators of the cortex-HR relationship. This study, conducted using a different fMRI pulse sequence, found comparable results in a slightly more dorsal area in the anterior portion of the anterior mid-cingulate cortex, which we refer to as the rostral dorsal cingulate (rdACC). Results in both studies were statistically quite strong and survived correction for multiple comparisons in multiple effects (Path a, Path b, and the mediation a*b effect).

**Medial prefrontal cortex, emotion, and autonomic control**

A wealth of non-human animal literature supports a connection between MPFC and autonomic control of the heart (Bandler et al., 2000; Barbas et al., 2003; Devinsky,
Morrell, & Vogt, 1995; Saper, 2002), and MPFC has been referred to as “visceromotor cortex,” in contrast to lateral orbital “viscerosensory” areas (Price, 1999). Interestingly, vmPFC has been linked to higher-order contextual control over stress responses as well, providing a conceptual link with the cognitive generation and regulation of stress responses in humans. For example, a recent series of studies has shown that rat vmPFC inactivation abolishes the beneficial effects of stressor controllability on fear responses (Amat et al., 2005), whereas vmPFC activation mimics the effects of control (Amat, Paul, Watkins, & Maier, 2008). vmPFC inactivation also blocked the “immunizing” effects of prior exposure (Amat, Paul, Zarza, Watkins, & Maier, 2006) on stressor reactivity. Another series of experiments has shown that inactivation of the vmPFC or hippocampus prevents consolidation of fear extinction (Corcoran & Quirk, 2007b; Sierra-Mercado, Corcoran, Lebron-Milad, & Quirk, 2006), which is considered to be another type of safety-related context learning (Davis, 1992). Conversely, vmPFC stimulation potentiates or mimics extinction memory (Milad & Quirk, 2002; Milad, Vidal-Gonzalez, & Quirk, 2004).

Recent neuroanatomical, lesion, and electrophysiological evidence, however, also supports a functional distinction between dorsal and ventral MPFC sub-regions (Quirk & Beer, 2006). A potential rat homologue of the pgACC/mOFC distinction is the difference between the dorsal pre-limbic (PL) and ventral infralimbic (IL) medial frontal cortices. These two structures project to different subcortical nuclei (Gabbott, Warner, Jays, Salway, & Busby, 2005; Hoover & Vertes, 2007; McDonald, Mascagni, & Guo, 1996; Vertes, 2004)—the PL to the basolateral amygdala, associated with fear learning, and the IL to intercalated inhibitory neurons in the amygdala (Vertes, 2004), which are associated with the suppression of fear behavior during extinction. It is IL in particular in which stimulation potentiates fear extinction and lesions reduce it (Milad & Quirk, 2002; Quirk, Russo, Barron, & Lebron, 2000). In addition, stress induces dendritic retraction in IL (but not PL), an effect associated with impaired fear extinction (Izquierdo, Wellman, & Holmes, 2006). IL lesions also reduce stress-induced activity in hypothalamic pre-autonomic neurons, whereas PL lesions do not (Radley, Arias, & Sawchenko, 2006). Instead, PL activity is related to the expression of learned fear responses (Corcoran & Quirk, 2007a). All of the abovementioned information is consistent with the view that PL
activity promotes autonomic responses to stress, perhaps by mediating cognition and memory-related processes (Gabbott, Warner, Jays, & Bacon, 2003; Vertes, 2006), whereas IL activity inhibits them. The general pattern across these results matches the pattern of reciprocal control of HR in our study.

The pattern of reciprocal dorsal and ventral predictors of HR (though “dorsal” in this case is still relatively ventral, in the pgACC) replicates and extends findings from earlier human neuroimaging papers of stressful tasks. For example, in a large study with nearly 100 participants, Gianaros et al. (P. Gianaros et al., 2004) found positive and negative correlations with HR\(^1\) in the dorsal MPFC and vmPFC/mOFC, respectively. Critchley et al. (H. D. Critchley et al., 2000), in an important early study, investigated correlations between brain activity and HR and blood pressure increases with mental arithmetic stress and hand exercise. Increases during the two "stressor" tasks were found in the dorsal cingulate, and increases were positively correlated with blood pressure. Conversely, OFC showed decreases with the stressor tasks, and mOFC in particular was negatively correlated with blood pressure. In addition, though we did not perform tests of lateralization, our findings of right-sided effects in mOFC are consistent with work showing right-lateralized increases in HR reactivity with inactivation (with sodium amytal) (Ahern et al., 2001) and VMPFC damage (Hilz et al., 2006). The current findings extend this literature by showing this reciprocal relationship between dorsal and ventral subregions of the mPFC in response to social threat, as well as showing that although there is a functional coupling between these two regions, their effects on HR reactivity are independent of each other.

**Reciprocal medial frontal subregions and cardiac control**

The notion of reciprocal regulation of physiological threat responses by different medial prefrontal sub-regions has implications for models of vagal control and peripheral regulation, as well as parallels in both neuroanatomical and functional studies of emotion. One emerging idea is that the dorsal and ventral (medial orbital) MPFC are important for

---

\(^1\) Heart period, 1/HR, was actually analyzed, which has a more nearly linear relationship with underlying sympathetic and parasympathetic effectors (Cacioppo et al., 1994), but we discuss the results in terms of heart rate here for ease of interpretation.
higher brain control of the sympathetic and parasympathetic (vagal) control of the heart, respectively. Speech preparation stressors have been shown to robustly increase HR by both increasing noradrenergically mediated sympathetic output to the heart and reducing cholinergically mediated parasympathetic output, which normally inhibits HR (Berntson et al., 1994; Berntson, Quigley, & Lozano, 2007; Cacioppo et al., 1994; Thayer & Lane, 2007).

Recent neuroimaging studies in humans support the view that dorsal and ventral MPFC may be relatively specialized for sympathetic and parasympathetic control, respectively. In the Gianaros et al. study (P. Gianaros et al., 2004), decreases in mOFC were specifically correlated with reduced high-frequency heart-rate variability (HRV), a relatively pure measure of parasympathetic function, during performance stress. Similar findings by Lane et al. (2001) of positive associations between HRV and activity in the VMPFC support the parasympathetic-mOFC association, but a more recent study found positive correlations between HRV and both pgACC and rdACC (R. D. Lane et al., 2009). In this study, smaller HRV decreases elicited by emotional stimuli were correlated with larger pgACC responses.

Our measures of in-scanner heart period did not have the temporal precision to reliably measure more pure indices of parasympathetic effectors, such as HRV, nor did we use impedance cardiography to measure pre-ejection period, a more reliable index of sympathetic effects on the heart (Berntson et al., 2007). Also, our time series analyses precluded averaging over the relatively long time periods needed to obtain these measures. However, some clues can be gained by our analyses of timing. The sympathetic system affects HR with an approximately 2 sec longer response latency than the parasympathetic system, which can exert effects within hundreds of milliseconds (Berntson et al., 2007). In our study, the shorter latency of the mOFC-HR correlation relative to the dorsal MPFC and other regions supports the notion that it provides largely parasympathetic innervation of the heart, and that decreases in activity in this area reflect decreased parasympathetic outflow which increases HR.

Another point of convergence is that the combination of significant negative coupling but independent prediction of HR in pgACC and mOFC parallels other findings of negatively coupled but independent activity of sympathetic and parasympathetic
cardiac output. For example, Cacioppo et al. (Cacioppo et al., 1994), using a careful series of pharmacological blockades, provided evidence that the sympathetic and parasympathetic systems operated in a largely independent manner. Whereas we did find significant functional connectivity between pgACC (putatively primarily sympathetic) and mOFC (putatively parasympathetic), each of these regions made statistically separable contributions to predictions of HR rise and fall across time. In addition, these regions were statistically separable predictors of individual differences in HR reactivity.

**HR reactivity as an individual differences measure**

Our results also have implications for the characterization of individual differences in threat and stress reactivity. In our study, the degree to which an individual expressed each of the links between SET, brain activity, and HR strongly predicted the degree of HR reactivity to the challenge. HR reactivity provides a robust way of characterizing individual differences in physiological responses to stressors, rather than relying solely on subjective reports as outcome measures, which can depend on self-presentation (Weinberger, Schwartz, & Davidson, 1979) and generally have more complex determinants. HR reactivity as an individual difference measure has been associated with increased hypothalamic-pituitary-adrenal axis activity (al'Absi et al., 2000; Cohen et al., 2000; Sgoutas-Emch et al., 1994; Uchino et al., 1995) and cellular immune responses to acute psychological stressors (reviewed in (Cacioppo, 1994))(Knapp et al., 1992; Sgoutas-Emch et al., 1994; Uchino et al., 1995). HR reactivity to acute laboratory stressors in particular has been associated with blastogenic responses to mitogen (a probe of immune cell proliferation in response to challenge), NK cell counts, and lymphocyte cell counts (t-cells, b-cells) (Brosschot et al., 1992; Cohen et al., 2000; Knapp et al., 1992; Landmann et al., 1984; Sgoutas-Emch et al., 1994; Uchino et al., 1995). HR reactivity is an outcome deserving of study in its own right, and is not likely to be highly related to reported subjective anxiety or other psychological reports. In the present study, for example, HR reactivity was uncorrelated with emotional resilience, optimism, and other measures.

**Implications for cortical regulation of affective states**
It has long been recognized that autonomic physiology is an important aspect of emotional behavior (Darwin, 1904; Levenson, 2003). Some researchers have suggested that a main function of emotions is to coordinate behavioral and physiological responses to environmental demands (particularly threat), and others have claimed that autonomic activity even plays a role in shaping central social and emotional experience. In addition to its part in autonomic regulation, the MPFC plays a prominent role in “self-related” processes (Northoff et al., 2006) and emotions (T. Wager et al., 2008). It seems plausible, then, that the brain systems responsible for generating HR responses to threat, including reciprocal dorsal and ventral MPFC subsystems, may be important in the generation of valenced emotional states and their physiological effects more broadly. The association between MPFC, the cognitive generation and regulation of threat and safety appraisals, and autonomic control may underlie the widespread association of measures related to high parasympathetic tone (such as heart-rate variability) and diverse health benefits (Thayer & Lane, 2007; Thayer & Sternberg, 2006).

In a recent series of meta-analyses of 163 neuroimaging studies of emotion (Kober et al., 2008; T. Wager et al., 2008), we analyzed which brain areas were consistently associated with states of valenced emotional experience (as opposed to perceptions of “affective” stimuli such as facial expressions). Negative emotional experience was associated with increases in dmPFC, and positive emotional experience was associated with increases in vmPFC and mOFC (Wager et al., 2008; see (Etkin & Wager, in press), Figure 1). (Recall that vmPFC/mOFC was deactivated during SET, which elicits negative emotion, in the present study). Dorsal MPFC and pgACC emerged as distinct regions, both of which were co-activated across studies with activity around the peri-aqueductal gray (PAG), a key brainstem coordinator of autonomic responses to threat (Bandler & Shipley, 1994; Behbehani, 1995). A companion study addresses the role of the PAG in human SET responses (T. D. Wager et al., submitted).

Paralleling these results, individual studies have also shown evidence for differential functional roles of dorsal and medial PFC in emotion-induction tasks. For example, Porro et al. (Porro, 2003) reported increases in dorsomedial and decreases in ventromedial frontal activity associated with heart rate increases during pain anticipation. In another recent study, we found that rdACC activity was correlated with increased
amygdala activity and less successful modulation of emotional responses to negative pictures (T. D. Wager et al., 2008). Activity in the rostral MPFC (anterior to pgACC), however, was associated with nucleus accumbens increases and successful modulation of reported negative emotion. Rostral vmPFC activity has also been correlated with reduced cortisol responses to a combined speech-and-math stressor (Kern et al., 2008), reduced amygdala activity (Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007; Urry et al., 2006), and healthier (deeper) diurnal cortisol rhythms (Urry et al., 2006).

Another point of converging evidence for the positive/negative functional distinction between ventral and dorsal MPFC comes from studies of fear conditioning, which has been relatively well studied in human and animal models. In humans, several studies have found increases in the dorsal MPFC when conditioned stimuli (cues associated with shock) are presented, and reductions in mOFC (Delgado, Nearing, Ledoux, & Phelps, 2008; Phelps, Delgado, Nearing, & LeDoux, 2004a; Schiller, Levy, Niv, LeDoux, & Phelps, 2008). Animal fear conditioning studies also provide evidence supporting the dorsal/ventral MPFC distinction. For example, PL shows increases electrophysiological responses during trace conditioning to aversive cues, whereas activity in the IL cortex decreases during fear conditioning (Gilmartin & Mcechron, 2005).

Overall, there are two ways in which studying brain-heart relationships can be useful for understanding emotion and brain organization more broadly. As we discuss above, one is to characterize the central brain mechanisms and brain-body pathways that are likely to be implicated in a number of mental disorders. A second use is to promote the careful measurement and analysis of autonomic effects in other studies of emotional and cognitive function.

It is possible that brain activity thought to be related to other cognitive processes is actually related to autonomic control associated with cognition. For example, response inhibition, a basic cognitive control operation, is thought to induce transient parasympathetically mediated heart-rate slowing (Jennings, 1992). In addition, the vmPFC has been associated with top-down perceptual set effects (Summerfield et al., 2006) and long-term memory retrieval (summarized in (van Snellenberg & Wager, in press); (Buckner & Carroll, 2007)). In one study, we found that relative pgACC
increases during task switching was the best and only predictor of individual differences in the ability to shift attention rapidly across four separate task variants (T. D. Wager, Jonides, Smith, & Nichols, 2005). It could be that the brain processes that regulate physiological resource allocation during demanding task performance (i.e., MPFC) are key limiting resources on individuals’ performance capacity.

Absence of amygdala increases in SET: Differentiation of SET from other negative emotional tasks. Though reciprocal dorsal and ventral MPFC activity has emerged as a common theme across manipulations of emotion, the present results also provide evidence that not all aversive emotional states and outcomes are the same. For example, the basolateral and central amygdala have been strongly implicated in the learning and expression of cue-threat associations, respectively (LeDoux, 1996). Amygdala activity is a prominent correlate of anticipatory anxiety (Nitschke et al., 2009), conditioned fear in anticipation of shock (Phelps et al., 2004b), and reported negative emotion in response to affective pictures (Ochsner & Gross, 2008; Ochsner et al., 2004; Phan et al., 2005). Amygdala hyperactivity is a prominent feature of multiple anxiety disorders (Etkin & Wager, 2007), and its neurons encode both positive and negative predicted value (Paton, Belova, Morrison, & Salzman, 2006).

However, our speech preparation task evoked no detectable increases in the amygdala—rather, they showed evidence for decreases during SET that did not strongly predict HR. Flexible change-point models that could have detected activation that was uncoupled from the task demand and HR also showed no evidence for amygdala responses during the speech preparation period (Martin A. Lindquist et al., 2007). The lack of amygdala activation is consistent with other studies of performance stressors, some of which have produced reliable decreases in the amygdala and hippocampus in both PET and fMRI (Dedovic et al., 2005; Pruessner et al., 2008). Physical pain shows a similar pattern of amygdala unresponsiveness or decreases (Derbyshire et al., 1997; Petrovic, Carlsson, Petersson, Hansson, & Ingvar, 2004), though results vary across studies. This pattern contrasts with recent reports of increased amygdala responses in cardiovascular responders performing a stressful cognitive task (Gianaros et al., 2008a), perhaps due to differences in the nature of the stressor; future research should address this
issue. The latter study notwithstanding, the brain data in this study support the notion that SET responses are qualitatively distinct from some other forms of affective processing.

**Desiderata on the single-epoch and path modeling approaches**

The paucity of neuroimaging studies of SET is perhaps due to the fact that studying SET involves several unique challenges, which require innovative approaches to the design and analysis of neuroimaging studies. We describe these, and our approach, briefly below.

A key, relatively novel aspect of our task design was that we used a single-epoch approach to eliciting a threat state over a brief (2 min) period. This is an atypical fMRI task design, as the vast majority of blood oxygen level dependent (BOLD) fMRI studies have studied emotional responses by using briefly presented affective stimuli, or by contrasting emotional manipulation or performance stress with a control condition in alternating blocks (typically every 20-30 sec). The use of a single-epoch design was important because it is likely that social threat responses cannot be reliably turned on and off multiple times during scanning: Participants habituate to even a single repetition of a SET challenge (Berntson et al., 1994; Cohen et al., 2000; Kelsey et al., 1999). Experimental designs with prolonged challenges (i.e., low temporal frequency) more closely mirror emotion induction procedures in non-imaging settings (Fredrickson & Levenson, 1998; Fredrickson, Mancuso, Branigan, & Tugade, 2000) and are more likely to induce strong changes in emotional states. Indeed, the notion that emotional states evolve more gradually over time has motivated the use of novel fMRI techniques (Wang et al., 2005) and positron emission tomography (Kern et al., 2008; Phan, Wager, Taylor, & Liberzon, 2002; Pruessner et al., 2008) in the study of stress and emotion.

An fMRI study by Preussner et al. (Pruessner et al., 2008) that compared alternating blocks of math performance under stressful and non-stressful control conditions (the “Montreal Neuroimaging Stress Test”) illustrates this difficulty. The stress vs. control comparison revealed increased fMRI signal in areas associated with math performance and controlled response selection—premotor cortex, caudal dorsal cingulate, and occipital association areas—but none of the areas associated with emotional experience in humans (T. Wager et al., 2008) or stress generation and
modulation in animals. The use of single-epoch, low-frequency designs is not completely novel (Breiter et al., 1997; Eisenberger, Lieberman, & Williams, 2003), but demonstrating that standard BOLD fMRI can capture responses to a single-epoch SET challenge suggests that this may be a promising way to characterize brain systems that perceive and respond to social threat.

Another key choice in this paper was the choice to avoid the standard method of making inferences about psychological mechanisms based solely on subtraction methods and logic. The subtraction method has been used to compare stressful performance with non-stressful task performance (Gianaros, May, Siegle, & Jennings, 2005; Kern et al., 2008; Pruessner et al., 2008). However, it is difficult to conceive of control conditions that isolate the essential affective experience component of SET responses that is likely to lead to physiological changes. For example, Preussner et al. compared stressful mental arithmetic, which includes time deadlines, negative performance feedback, and expectations about normative performance, with a control condition without these elements. It is difficult to determine which brain correlates of SET are related to aspects of the emotional state that generate subjective feelings and shape autonomic activity, and which reflect changes in how arithmetic operations are performed.

A complete parsing of brain activity induced by “stressor” tasks into cognitive, emotional, motoric, and autonomic afferent and efferent components is still likely to be a long way off. For example, our results do not inform on whether MPFC activity is linked to physiology because it is the seat of subjective emotional experience or some physiological mechanism that shapes homeostatic and metabolic processes largely outside of conscious awareness. Rather than attempting to isolate subjective experiences of emotion, we attempted to localize close correlates of integrated autonomic output. This output is of interest in its own right because of its consequences for the body, but it may also aid in the study of emotional experience, if only by identifying a potential alternative explanation for “emotional” as well as “cognitive” activations in the VMPFC.

*Challenges with physiology-induced imaging artifacts*

One potential problem with any study relating brain activity to HR, blood pressure, or emotional states that induce physiological changes is the possibility that the
psychological demands might induce vascular changes and thus fluctuation in fMRI signal that are non-neuronal in origin. There are several lines of defense against interpreting the results of the present study as vascular artifacts.

First, the pattern of results across the brain is very specifically localized to regions known to be involved in autonomic control from converging lesion, electrophysiological and neuroanatomical studies in animals and humans. Second, that argument notwithstanding, one way to assess whether this pattern is artifactual is to examine whether studies of mechanical changes in vascular responses produce similar results. The evidence to date suggests that they do not. For instance, hypercapnia (increased blood CO2 induced, for example, by breath-holding) results in increases in fMRI signal throughout the brain, rather than in specific regions (Thomason, Burrows, Gabrieli, & Glover, 2005; Vazquez et al., 2006). The pgACC and mOFC show relatively low changes in percent signal during a breath-hold challenge relative to other regions of the brain (see Thomasen et al., Figure 6). Third, many kinds of physiological artifacts, such as effects of pulsatile motion due to the heart-beat, occur at different temporal frequencies from neuronal-induced fMRI signal, and are not likely to be significant sources of the brain-HR covariance observed here (see Birn et al. this issue). Finally, in our second, companion study (Wager et al., this issue) we were able to identify and control for signal in major arteries, including the anterior cerebral artery at the genu of the corpus callosum, and the MPFC findings were replicated.

Visceromotor or viscerosensory? Causality and limitations in inference from path models. It is important to consider which functional connections can reliably be given a causal interpretation and which cannot. In the present study, SET was experimentally manipulated in an off-on-off design (ruling out confounds related to task order), and we suggest that resulting brain activity can be interpreted as caused by SET. Brain relationships with HR could plausibly be either causes of HR change, effects of HR as perceived by the brain, or both; however, the locations of active regions in “visceromotor” cortex (Price, 1999) favor brain-to-heart causality. However, this causal interpretation rests on converging evidence, rather than on the data or analyses from this study. Causal inferences about the direction of connectivity between two observed (not
manipulated) variables is problematic even in extremely large samples, and causal inferences from formal causal models require a number of assumptions that are difficult to evaluate and seldom hold (Rubin, 1997; Sobel, 1995). Thus, we interpret our results as evidence for limited effective connectivity (functional relationships that remain after controlling for a number of known confounding sources), rather than evidence for causal relationships.

A note on the relative timing of brain-HR relationships is relevant as well. The analyses we present here suggest that relationships between HR and observed fMRI activity are roughly simultaneous, with the exception of a shorter functional latency between the mOFC and HR. However, neural activity occurs on average approximately 5-6 seconds before the peak of the observed BOLD response (Aguirre et al., 1998; T. D. Wager, Vazquez, Hernandez, & Noll, 2005), implying that neural activity preceded HR activity in all of these regions. We do not emphasize this latency as a source of causal information primarily because the HR quantification contains lower-frequency information that may delay apparent HR changes. The quantification of relative brain-heart latencies is a promising avenue for future exploration.

**Conclusion**

In conclusion, these findings contribute significantly to the investigation of brain-physiology relationships in the context of social threat. The relationship between social threat and associated physiological responses (HR) was mediated by reliable and sustained increases in pre-genual cingulate/MPFC and decreases in vmPFC/mOFC. Future investigations should more specifically examine whether the dorsal/ventral distinction within the MPFC maps onto distinctions within the peripheral nervous system (sympathetic and parasympathetic, respectively), their relationships to emotion reports (negative and positive, respectively), and their potential subcortical mediators. We address some of these issues in a companion paper (T. D. Wager et al., submitted).
References


Serial pathways from primate prefrontal cortex to autonomic areas may influence emotional expression, 4 Cong. Rec. 25 (2003).


When the social self is threatened: shame, physiology, and health, 72 Cong. Rec. 1191-1216 (2004).


CORTICAL MEDIATORS OF SOCIAL THREAT


... stress-induced abnormal left ventricular function response in patients with coronary artery disease ... (2001).


Table 1.

<table>
<thead>
<tr>
<th>Group</th>
<th>Name</th>
<th>1st-level: within-subjects</th>
<th>2nd-level: moderation by average HR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>Increases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial frontal</td>
<td>pgACC</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Temporal</td>
<td>vTC</td>
<td>25</td>
<td>-56</td>
</tr>
<tr>
<td></td>
<td>STS</td>
<td>-66</td>
<td>-41</td>
</tr>
<tr>
<td>Occipital</td>
<td>OCC</td>
<td>3</td>
<td>-88</td>
</tr>
<tr>
<td></td>
<td>R OCC</td>
<td>22</td>
<td>-91</td>
</tr>
<tr>
<td>Decreases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orbitofrontal</td>
<td>L mOFC</td>
<td>-19</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>R mOFC</td>
<td>19</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>R aOFC</td>
<td>22</td>
<td>59</td>
</tr>
<tr>
<td>Basal ganglia</td>
<td>R Putamen</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>Insula</td>
<td>R Ins</td>
<td>44</td>
<td>9</td>
</tr>
<tr>
<td>Temporal</td>
<td>R TP</td>
<td>38</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>R ITG</td>
<td>86</td>
<td>-19</td>
</tr>
<tr>
<td></td>
<td>R STG</td>
<td>59</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>L STG</td>
<td>-56</td>
<td>-3</td>
</tr>
<tr>
<td>Diencephalon</td>
<td>Thal</td>
<td>9</td>
<td>-25</td>
</tr>
<tr>
<td></td>
<td>R SPL</td>
<td>44</td>
<td>-86</td>
</tr>
<tr>
<td>Posterior medial</td>
<td>PCC</td>
<td>0</td>
<td>-53</td>
</tr>
<tr>
<td></td>
<td>MidCing</td>
<td>3</td>
<td>-16</td>
</tr>
<tr>
<td></td>
<td>Precun</td>
<td>-9</td>
<td>-59</td>
</tr>
<tr>
<td></td>
<td>Precun</td>
<td>3</td>
<td>-50</td>
</tr>
<tr>
<td>Sensorimotor</td>
<td>SMC</td>
<td>6</td>
<td>-28</td>
</tr>
<tr>
<td></td>
<td>SMC</td>
<td>-6</td>
<td>-31</td>
</tr>
</tbody>
</table>
Table 2.

<table>
<thead>
<tr>
<th>Positive association</th>
<th>1st-level: within-subjects</th>
<th>2nd-level: moderation by average HR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
<td><strong>Name</strong></td>
<td><strong>x</strong></td>
</tr>
<tr>
<td><strong>Medial frontal</strong></td>
<td>pgACC</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>vmPFC</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>rdACC</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>rMPFC</td>
<td>-12</td>
</tr>
<tr>
<td></td>
<td>preSMA</td>
<td></td>
</tr>
<tr>
<td><strong>Lateral frontal</strong></td>
<td>L DLPFC</td>
<td>-25</td>
</tr>
<tr>
<td></td>
<td>L SFG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R DLPFC</td>
<td></td>
</tr>
<tr>
<td><strong>Basal ganglia</strong></td>
<td>R Cau</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>R Cau</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>L Cau</td>
<td>-16</td>
</tr>
<tr>
<td></td>
<td>R Cau</td>
<td>19</td>
</tr>
<tr>
<td><strong>Negative association</strong></td>
<td>R mOFC</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>R mOFC</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>R IOFC</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>R aOFC</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>L IOFC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L aOFC</td>
<td></td>
</tr>
<tr>
<td><strong>Medial frontal</strong></td>
<td>VMPFC/mOFC</td>
<td></td>
</tr>
<tr>
<td><strong>Posterior medial</strong></td>
<td>MidCing</td>
<td>3</td>
</tr>
<tr>
<td><strong>Lateral frontal</strong></td>
<td>R IFG</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>R IFG</td>
<td>44</td>
</tr>
<tr>
<td><strong>Temporal</strong></td>
<td>L TP</td>
<td>-28</td>
</tr>
<tr>
<td></td>
<td>L TP</td>
<td>-41</td>
</tr>
<tr>
<td></td>
<td>L STG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R TC/OCC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L PHCMP</td>
<td>-31</td>
</tr>
<tr>
<td><strong>Basal ganglia</strong></td>
<td>R Putamen/vINS</td>
<td>38</td>
</tr>
<tr>
<td><strong>Insula</strong></td>
<td>R valINS</td>
<td>44</td>
</tr>
</tbody>
</table>
Table 3.

<table>
<thead>
<tr>
<th>Amygdala region</th>
<th>Path a (SET)</th>
<th>Path b (HR)</th>
<th>a*b</th>
<th>Path a (SET)</th>
<th>Path b (HR)</th>
<th>a*b</th>
</tr>
</thead>
<tbody>
<tr>
<td>L corticomedial</td>
<td>-0.26* (0.11)</td>
<td>-0.03 (0.03)</td>
<td>0.02 (0.01)</td>
<td>-0.11 (0.09)</td>
<td>-0.04 (0.03)</td>
<td>0.02 (0.01)</td>
</tr>
<tr>
<td>R corticomedial</td>
<td>-0.39* (0.15)</td>
<td>-0.06 (0.04)</td>
<td>0.05** (0.02)</td>
<td>-0.39 (0.12)</td>
<td>-0.06+ (0.03)</td>
<td>0.07** (0.02)</td>
</tr>
<tr>
<td>L basolateral</td>
<td>-0.18+ (0.09)</td>
<td>-0.02 (0.03)</td>
<td>0.01 (0.01)</td>
<td>-0.06 (0.06)</td>
<td>-0.02 (0.02)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>R basolateral</td>
<td>-0.26* (0.10)</td>
<td>-0.15** (0.04)</td>
<td>0.06+ (0.03)</td>
<td>-0.11 (0.08)</td>
<td>-0.13** (0.03)</td>
<td>0.05+ (0.03)</td>
</tr>
</tbody>
</table>

Note. ROI in anatomically defined amygdala subregions. Negative effects at the first level indicate de-activation. Negative effects at the second level indicate more negative responses in high HR reactors. +, p < .10; *, p < .05; **, p < .01.
Table 4.

<table>
<thead>
<tr>
<th>Overlap of 1st-level SET and HR (a and b paths)</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Vol. (mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive a and b paths</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial frontal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pgACC</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>5420</td>
</tr>
<tr>
<td>R. Cau</td>
<td>12</td>
<td>16</td>
<td>9</td>
<td>29</td>
</tr>
<tr>
<td>Negative a and b paths</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orbitofrontal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. mOFC</td>
<td>16</td>
<td>41</td>
<td>-18</td>
<td>117</td>
</tr>
<tr>
<td>R. aOFC</td>
<td>19</td>
<td>59</td>
<td>-18</td>
<td>59</td>
</tr>
<tr>
<td>Basal ganglia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. Putamen</td>
<td>34</td>
<td>0</td>
<td>0</td>
<td>59</td>
</tr>
</tbody>
</table>
Table 5.

### Multi-level path model results

<table>
<thead>
<tr>
<th></th>
<th>pgACC</th>
<th></th>
<th>R medial OFC</th>
<th></th>
<th>R Putamen</th>
<th></th>
<th>Direct</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SET</td>
<td>HR</td>
<td>Mediation</td>
<td>SET</td>
<td>HR</td>
<td>Mediation</td>
<td>SET</td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td>Path a</td>
<td>Path b</td>
<td>Path a*b</td>
<td>Path a</td>
<td>Path b</td>
<td>Path a*b</td>
<td>Path a</td>
</tr>
<tr>
<td>STE</td>
<td></td>
<td>1.34</td>
<td>0.11</td>
<td>0.19</td>
<td>-1.12</td>
<td>-0.07</td>
<td>0.12</td>
<td>-0.71</td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>0.20</td>
<td>0.02</td>
<td>0.04</td>
<td>0.19</td>
<td>0.02</td>
<td>0.04</td>
<td>0.14</td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>0.0001</td>
<td>&lt;.0001</td>
<td>0.0029</td>
<td>0.0045</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

### Second Level Moderation by average HR response to SET

|            |          | SET      | HR           | Mediation| SET       | HR       | Mediation| SET   | HR    | Mediation| Direct | Total   |        |
|------------|----------|----------|--------------|----------|-----------|----------|----------|--------|--------|----------|--------|---------|        |
| Slope      | 0.78     | 0.07     | 0.18         | -0.55    | -0.06     | 0.10     | -0.43    | -0.02  | 0.04   | 0.40     | 0.76   |        |        |
| STE        | 0.16     | 0.02     | 0.03         | 0.16     | 0.02      | 0.03     | 0.11     | 0.01   | 0.02   | 0.06     | 0.05   |        |        |
| t          | 4.74     | 4.00     | 5.25         | -3.43    | -3.12     | 3.33     | -3.97    | -1.40  | 2.33   | 6.84     | 13.96  |        |        |
| p          | 0.0001   | 0.0006   | <.0001       | 0.0026   | 0.0053    | 0.0032   | 0.0008   | 0.1760 | 0.0298 | <.0001   | <.0001 |        |        |

Note. Multi-level path model results. Results correspond to the path model diagrammed in Figure 6. STE: standard error.
Figure Captions

Figure 1. A) Task design. The SET manipulation involved a 2-min resting baseline, a 15 sec visual presentation of the speech topic, a 2-min preparation period, a 15-sec no speech instruction, and a 2.5 min recovery period. B) Heart rate changes over time. Individuals are shown by the light gray lines, and the group average (with shaded standard error region) is shown by the heavy black line. C) Nuisance regressors for fMRI analysis. These varied across participants, but always included regressors for linear and higher-order head movement, potential outliers based on task- and physiology-blind global signal analysis, global activity values, and linear drift.

Figure 2. Path model (top) and results of the multi-level mediation effect parametric mapping search. Relationships between Speech Prep and brain activity (Path a) and between fMRI activity and heart rate (Path b, controlling for the Speech Prep regressor) were tested both within person and between persons, with heart rate (HR) reactivity as a predictor of individual differences in path amplitude. A) Path a results for the first-level model (time series). Sagittal slice showing regions whose activity increased (yellow/orange) or decreased (blue) in response to the social evaluative threat (SET) challenge. Significant regions of 3 or more contiguous voxels at q < .05 False Discovery Rate corrected, and contiguous regions at p < .005, are shown. B) Path b results for the first-level model (time series). Sagittal slices showing significant positive (yellow/orange) or negative (blue) correlates of heart rate changes over time, controlling for Speech Prep regressor. C) Path a results for the second-level model, showing correlations between SET-brain connectivity and HR reactivity. Positive correlations are shown in yellow, and negative correlation in blue. D) Path b results for the second-level model, showing correlations between brain-HR connectivity and HR reactivity.

Figure 3. Regions showing both Path a (SET responses) and Path b (prediction of HR) in the first-level (time series) model, and results from one amygdala region of interest (ROI). Positive results for both Path a and Path b are shown in red, and negative results for both are shown in blue. No regions showed positive a and negative b effects or vice versa. Results are shown at P < .001 for display, but all regions showed significant effects in both paths at q < .05 FDR (p <
CORTICAL MEDIATORS OF SOCIAL THREAT

The time series plots at right show group-averaged time series data across the run estimated with the Hierarchical Exponentially Weighted Moving Average (HEWMA) model. They are similar to, but smoother, than the raw data averages. Instruction periods are shown as yellow horizontal bars, and the Speech Prep period is shown as a blue horizontal bar in each plot. The HEWMA model provides estimates of which time points are deviant from the pre-SET baseline using a 2-state mixture model; these periods are marked with a red line. Time points that were individually significantly different from the average pre-scan baseline (zero on the y-axis) are marked with green dots at y = 0. The right (R) putamen alone showed evidence for a trend towards de-activation even before the speech instruction onset, as evidenced by a change-point value that occurred before the instruction. The right amygdala ROI showed significant de-activation in response to SET, but this activity did not predict heart rate fluctuations. Other amygdala sub-regions showed similar results.

**Figure 4.** Visualization of brain-HR connectivity (related to Path b) for the pregenual cingulate. A) Superimposed plots of the group-average time series data (black, with standard error regions shaded) against the group-average HR data (red). Speech Prep-related variance was not removed for display purposes, so the response to the social threat challenge can be seen. In addition, nuisance covariates (see Figure 1C) were not removed for display purposes, so the original response in both brain and heart can be seen. Only standard preprocessing procedures were performed. Instruction periods are shown as yellow horizontal bars, and the Speech Prep period is shown as a blue horizontal bar. B) The same plot as in (A), but with nuisance covariates removed. C) Plots of brain (black) and superimposed HR (red) for individual subjects, each shown in a separate panel. D) Correlation values between subjects’ brain activity time series and their own HR time series vs. the group-average HR time series.

**Figure 5.** Path diagram and effect plots for the pregenual anterior cingulate (pgACC). A) Path a results for Level 1 (time series SET-brain relationship) and Level 2 (correlation between SET-brain relationship and HR reactivity). Mean path coefficients are shown, with standard errors in parentheses. ***, P < .001. The line plot (left panel) shows the first-level effects, the relationships between the SET predictor (which took on values of 0 for baseline and 1 during speech preparation; x-axis) and fMRI activity (y-axis). Relationships for Individual participants
are shown as blue lines, one per participant. The group-average effect with its standard error is shown by the black line and gray shaded area. The right panel shows a scatterplot of the second-level relationship between individual differences in the slope of the Path $a$ effect (x-axis) and the average HR response to the task (y-axis). The significant relationship ($\tau = .68, p < .00037$ [the FDR threshold]) indicates that those with high HR reactivity showed larger SET-brain (Path $a$) effects. B) The same relationships for the brain-HR relationship (Path $b$), controlling for the SET predictor. Significant first- and second-level effects demonstrate the reliable link between individual profiles of brain activity and individual profiles of HR changes across time.

**Figure 6.** Mediation path diagram for all three key mediators of social evaluative threat (SET) effects on heart rate. Solid black lines indicate significant relationships, and light gray lines indicate non-significant relationships. Connections hypothesized to be directional are shown as one-way arrows, whereas effects likely to be bi-directional (feedback loops) based on anatomy are shown as double-headed arrows. Causality could only be inferred for the SET-brain effects because SET was experimentally manipulated. First-level effects (SET-brain connectivity for Path $a$ or brain-HR connectivity for Path $b$) are shown as line plots, and second-level effects (correlations between Path $a$ or Path $b$ and HR reactivity) are shown as scatterplots. Effect plots are shown for pregenual cingulate (pgACC) and medial orbitofrontal cortex (mOFC), but are omitted for putamen (Put) for space reasons. Full statistics are presented in Table 5.

**Figure 7.** Bayesian estimates of cross-correlation latency between brain activity and heart rate for key regions. Mean latency across individuals is shown by a dot for each region, and standard errors across participants are shown by the black lines. Zero cross-correlation indicates that the best correlations were found with no relative lag of the brain and heart rate (HR) time series. Blue dots indicated that deactivation to SET predicted HR increases, and red dots indicated that activation predicted HR increases. We note that the estimation procedure does not account for an approximately 6 sec lag between neural activity in each region and the observed BOLD response. The bottom plots show the full mean cross-correlation estimates (with shaded standard errors) for the medial orbitofrontal cortex (mOFC) and the pregenual cingulate (pgACC). The mOFC correlation values are shifted somewhat towards negative values, indicating a brain
response that is lagged by the HR response. This shift was statistically significant when comparing the mOFC and pgACC lag values directly.
Figure 1.

A  Task design: Social Evaluative Threat (SET)

Speech topic instruction  Speech preparation  “No speech” instruction

Fixation baseline  Fixation baseline

Time (sec)

B  Heart-rate increases

Heart Rate (vs. baseline)

Time (sec)

C  Example covariates for one subject

Movement

Drift, Outliers

Time (TRs, 2 sec)
Figure 3. Overlap of 1st-level a and b effects

**Pregenual anterior cingulate**

**R Orbitofrontal cortex**

**R Putamen**

*a priori ROI: R amygdala*
CORTICAL MEDIATORS OF SOCIAL THREAT

Figure 5.
Multi-level Mediation: pgACC

Speech Prep (SET) → Path a → pgACC activity → Path b → Heart rate

Path a:
- Level 1: .72(.15)***
- Level 2: .52(.12)***

Path b:
- Level 1: .26(.05)***
- Level 2: .18(.04)***

A. Within-subjects b effect
- Brain activity (BOLD units) vs. SET predictor
- Average heart-rate increase vs. a path slope

B. Within-subjects a effect
- Heart rate vs. Brain activity
- Average heart-rate increase vs. b path slope

Level 2 (ind. diffs) with correlation coefficient r = 0.68 and r = 0.71.
CORTICAL MEDIATORS OF SOCIAL THREAT

Figure 6.

pgACC: SET

pgACC: Heart Rate

mOFC: SET

mOFC: Heart Rate

Brain mediators

Speech (SET)

Heart rate

pgACC

R mOFC

Put
Figure 7.