Multivariate Brain Prediction of Heart Rate and Skin Conductance Responses to Social Threat

Hedwig Eisenbarth,1,2 Luke J. Chang,1 and Tor D. Wager1

1Department of Psychology and Neuroscience and the Institute for Cognitive Science, University of Colorado, Boulder, Colorado 80309, 2Department of Psychology, University of Southampton, Southampton, Hampshire SO17 1BJ, United Kingdom, and 3Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755

Psychosocial stressors induce autonomic nervous system (ANS) responses in multiple body systems that are linked to health risks. Much work has focused on the common effects of stress, but ANS responses in different body systems are dissociable and may result from distinct patterns of cortical–subcortical interactions. Here, we used machine learning to develop multivariate patterns of fMRI activity predictive of heart rate (HR) and skin conductance level (SCL) responses during social threat in humans (N = 18). Overall, brain patterns predicted both HR and SCL in cross-validated analyses successfully (r_{HR} = 0.54, r_{SCL} = 0.58, both p < 0.0001). These patterns partly reflected central stress mechanisms common to both responses because each pattern predicted the other signal to some degree (r_{HR→SCL} = 0.21 and r_{SCL→HR} = 0.22, both p < 0.01), but they were largely physiological response specific. Both patterns included positive predictive weights in dorsal anterior cingulate and cerebellum and negative weights in ventromedial PFC and local pattern similarity analyses within these regions suggested that they encode common central stress mechanisms. However, the predictive maps and searchlight analysis suggested that the patterns predictive of HR and SCL were substantially different across most of the brain, including significant differences in ventromedial PFC, insula, lateral PFC, pre-SMA, and dmPFC. Overall, the results indicate that specific patterns of cerebral activity track threat-induced autonomic responses in specific body systems. Physiological measures of threat are not interchangeable, but rather reflect specific interactions among brain systems.

Key words: heart rate; machine learning; multivariate pattern analysis; skin conductance; social threat; stress

Significance Statement

We show that threat-induced increases in heart rate and skin conductance share some common representations in the brain, located mainly in the vmPFC, temporal and parahippocampal cortices, thalamus, and brainstem. However, despite these similarities, the brain patterns that predict these two autonomic responses are largely distinct. This evidence for largely output-measure-specific regulation of autonomic responses argues against a common system hypothesis and provides evidence that different autonomic measures reflect distinct, measurable patterns of cortical–subcortical interactions.

Introduction

Stressful experiences such as social threat lead to physiological changes that are often adaptive (Cannon, 1935; Kemeny, 2003). However, these changes have a pronounced impact on the brain and body that can be harmful if stress experiences are chronic or prolonged (Juster et al., 2010). Stress-related physiological changes can be elicited by threats to social and physical well-being (Campbell and Ehlert, 2012; Wager and Gianaros, 2014) and are thought to contribute to cardiovascular disease (Everson-Rose and Lewis, 2005), anxiety and depression (Uliazev et al., 2012), and other problems. Social threats are particularly pervasive in modern society, so the basis for their effects on physiology is important to understand. Although there are many studies of stressful situations, relatively few relate brain activity to physiology directly (Wager and Gianaros, 2014) and, surprisingly, no prior studies to our knowledge compare the brain correlates of multiple specific autonomic and neuroendocrine responses directly. Such comparisons are needed to move beyond “stress responses” in the brain as a unitary phenomenon and begin to link specific brain processes with specific physiological measures (Levenson, 2014).

One reason for this gap in the literature is that theories of stress have historically focused on commonalities in physiological responses because stressors are thought to drive responses in multiple body systems in a coordinated fashion (Cannon, 1935; Taylor and...
physiological responses in different somatic systems (e.g., responses in the heart, skin, and other organs) are likely mediated by different patterns of interactions among brain circuits (Jänig and McLachlan, 1992; Norman et al., 2014). Autonomic measures are modestly correlated at best, even under stress ($r = 0.42–0.54$ in Lazarus et al., 1963; $r = 0.36–0.60$ in Taylor and Epstein, 1967; $r = 0.16–0.37$ in Croft et al., 2004), and HR and SCL load on different factors in multivariate analyses (Cuthbert et al., 2000). This parallels findings of loosely coupled pattern generators in brainstem autonomic systems in nonhuman animals (Jänig and McLachlan, 1992; Saper, 2002).

Therefore, whether stress-related autonomic changes reflect activation of one coherent brain system or multiple patterned responses is unclear (Levenson, 2014). The answer to this question has important implications for health because responses in different organs are differentially related to health risks. For example, acute HR reactivity is a specific risk factor for cardiovascular disease (Lennings et al., 2004).

In our study, we address the need to compare the brain patterns related to different stress-related autonomic responses by analyzing brain activity related to moment-by-moment time courses of HR and SCL during social threat. We expected physiology-related patterns to be distributed across brain regions, requiring integration across multiple systems to capture the physiological response; therefore, we used machine learning to identify multivariate patterns optimized to predict each of HR and SCL (both within and across participants). We used cross-predictive accuracy (Woo et al., 2014; Chang et al., 2015; Osher et al., 2016), similarity in multivariate predictive maps (Woo et al., 2014), and local spatial similarity analyses (Chikazoe et al., 2014; Haynes, 2015) to identify both common and distinct cerebral correlates of HR and SCL. We expected to find common predictors of both responses in central “threat appraisal systems” detailed below.

Materials and Methods

Participants

Eighteen healthy, right-handed, native-English-speaking students (mean age 21 years, 9 males) were recruited at Columbia University. Exclusion criteria were a prior history of neurological or psychiatric illness or current or prior psychoactive medication. Participants were asked to abstain from tobacco and caffeine use for 24 h before scanning. All participants gave written informed consent and the study was approved by the Columbia University Institutional Review Board. This study is a reanalysis of existing data (Wager et al., 2009a), but presents novel analysis methods and conclusions.

Procedure and fMRI task design

Before scanning, participants were informed that during scanning they would be given 2 min periods to prepare two different speeches mentally (Fig. 1). The following instructions for these 2 speeches were given before participants entered the scanner: (1) speeches should be 7 min long and would be presented to 2 different audiences after the scanning phase: 1 speech would be given before a panel of professors and experts in the law and business and the second would be scored by a computer analysis program, latent semantic computer analysis (LSA), which is capable of grading college-level essays (Landauer et al., 1998) (pictures and biographies of panelists were shown); (2) the speech topics would be presented during fMRI scanning; and (3) for control purposes, there was a small chance that the prepared speech would not actually be given after the scanning phase (no participants actually gave speeches).

As shown in Figure 1, after anatomical scans and baseline physiological and brain data acquisition for 120 s, the first speech topic was presented for 15 s on the screen and participants then had 2 min to prepare their speech silently. After that, on-screen instructions (15 s) with the topic of the second speech were presented. Again, participants were given 2 min to prepare the second speech. After this, every participant was told that they were randomly selected to not give a speech and asked to relax for the remaining 2 min. This period was used to measure recovery and deconfound effects of the social threat task from effects of habituation and/or fatigue during scanning. The two speech topics were: “the effects of interest rates on stock prices” and “the relationship between tariffs and free trade.” These topics were selected based on data from a pilot study, which asked a separate group of participants to rate their anticipated anxiety to a number of possible topics (data not shown). Assignment of the two topics and the two potential audiences (professors vs LSA) to the first or second speech preparation period were counterbalanced across participants.

During baseline, speech preparation, and recovery, participants were visually cued every 20 s to provide a current subjective anxiety rating on a continuous visual analog scale using an MR-compatible trackball (Resonance Technologies) with the right hand, ranging from “no anxiety” to “extremely anxious.” These ratings were interpolated to the TR (2 s) using linear interpolation. All in-scanner stimuli were presented by a digital projection onto a screen placed in the scanner room. Stimulus presentation was controlled by E-prime software (Psychology Software Tools).

Data acquisition and analysis

HR was collected continuously during fMRI acquisition using photoplethysmography (finger pulse) on the left index finger with a sampling rate of 100 Hz. From those data, successive peaks were identified using a custom algorithm identifying deviations from a moving average baseline implemented in Matlab (The MathWorks, RRID:SCR_001622). The algorithm identified deviations from baseline photoplethysmography using the intensity and derivative of photoplethysmography waveforms after using moving average regression to adjust for baseline drift. This provided automated finger-pulse peak detection. The positions of the peaks were reviewed manually and corrected by a coder blinded to conditions using a systematic review process built into the software. HR time series were calculated by taking the derivative of the peak-to-peak intervals, which provides an estimate of instantaneous HR. The time series was downsampled to the TR (2 s) by averaging the HR within each TR, which also provides some smoothing to reduce potential high-frequency artifacts. Next, the entire time series was detrended linearly by regressing out linear
time effects using the detrend function in MATLAB (detrend.m; The MathWorks, RRID:SCR_001622). We did not perform any other high-pass filtering apart from detrending. Finally, the entire time series for each participant was transformed into z-scores to ensure that the predictions across each physiological signal and across individuals would be on the same scale.

Skin conductance data were collected with leads placed on the volar surfaces of the first and third fingers of the left hand. SCL data were acquired at 100 Hz, down-sampled to the TR by averaging within each TR, subjected to linear detrending, and z-scored within each participant.

MR images were collected on a 1.5 Sigma Twin Speed Excite HD scanner (GE Medical Systems). Structural images were acquired using high-resolution T1-spoiled gradient recall images for anatomical localization and warping to a standard space (the MNI152 average brain). Functional images were acquired with a T2*-sensitive EPI BOLD pulse sequence (TR = 2000 ms, TE = 40 ms, flip angle = 60°) sensitive to blood-oxygen-level-dependent (BOLD) magnetic susceptibility. For each participant, 266 brain volumes were obtained during the scanning run (24 ascending odd/even interleaved slices, 3.4375 × 3.4375 × 4.5 mm).

Functional images were subjected to standard preprocessing. First, slice timing acquisition correction and realignment of the functional images to correct for head movement were performed using FSL (FMrib Center, University of Oxford). Remaining preprocessing steps were performed using the Statistical Parametric Mapping analysis package (SPM2, Wellcome Department of Cognitive Neurology, London, UK, RRID:SCR_007037). The realigned images from each participant were coregistered to the anatomical space of the structural image. Structural images were normalized to the Montreal Neurological Institute (MNI) template space (avg152t1.img). Finally, the normalized functional images were smoothed with an 8 mm Gaussian smoothing kernel to facilitate intersubject registration in group analysis. The first 4 images (8 s) were excluded from analyses. To avoid adding any additional signal loss of the fMRI signal, we did not apply any high-pass (or low-pass) filtering to the data because our task design used a very low temporal frequency. We also did not detrend the fMRI data for the analyses presented here after additional analyses using detrended data showed no appreciable impact on the results.

**Predictive modeling analyses**

To develop multivariate brain patterns predictive of autonomic nervous system (ANS) signals, we trained a whole-brain pattern map to predict standardized HR and SCL separately using principal components regression (Wager et al., 2013; Chang et al., 2013; and see below for more details). We did this separately for each participant and made predictions both about new physiological data from individualized (idiographic) maps and about new individuals (between-subject, population-level maps).

One of the main advantages of using a multivariate approach over the standard univariate approach is that we can identify a unique configuration of weights distributed throughout the brain that contribute independently to the prediction of a physiological signal. This allows us to not only assess how much variance in each physiological signal can be accounted for by our multivariate brain representations, but also to examine which weights in the model contribute more to the prediction of one signal relative to the other.

The basic premise of our prediction methods is to identify a distributed multivariate representation of a physiological signal. To accomplish this, we use penalized principal components regression (Hastie et al., 2005) to predict a standardized physiological signal using a vectorized representation of all voxels in a 2 × 2 × 2 mm whole-brain mask. Because there were more features (n = 352,328 voxels) for a given participant than we had observations (n = 266 TRs), we performed a principal components analysis decomposition of the brain data (within the cross-validation loop; see below) to reduce the number of features in the model and to estimate regression coefficients on distributed, coherent component maps rather than voxels. This allows us to use a reduced number of components that retain the full variance of the data to predict the physiological time series. We use L1-regularized regression to estimate the model parameters (least absolute shrinkage and selection operator; LASSO). These estimated β weights on the components are then back projected into voxel space and then back into 3D space. This results in a 3D brain model that can be applied to new test brain data to return a scalar value reflecting the predicted physiological response for a given brain image using the dot-product method or spatial correlation (Wager et al., 2013; Chang et al., 2013).

Our methods were optimized to account for several methodological challenges (Fig. 2) such as auto-correlation of the continuous data (the selection of idiographic cross-validation below). To reduce the computational demands for large amounts of RAM (matrix size = 4788 volumes × 352,328 voxels), we used a two-level training approach. Patterns were trained separately for each participant (idiographic cross-validation) and then combined at the group level by taking the average weight for each voxel across participants (group-based cross-validation). Therefore, we used two different types of cross-validation to train and evaluate the models. All custom functions used for these analyses are publicly available on our GitHub repository (https://github.com/canlab/CanlabCore).

**Idiographic cross-validation.** To account for autocorrelation in predicting the HR/SCL time series, we used a cross-validation technique developed for training forecasting models called rolling hvg-block cross-validation, based on a suggested algorithm by Racine (2000). As shown in Figure 2A, the idea is to denote a test set of v images, with a buffer of images h collected just earlier and later from the training set so that influences of the training data on test data due to autocorrelation are minimized and independence of training and testing data are maintained. The training set is a set of images collected just before and after the h-image buffer with a fixed sample size to maintain a consistent number of images in terms of asymptotic convergence. The selection of training and test sets is “rolling”, in the sense that hvg blocks successive across time are constructed until every image is part of at least one test set. Here, we excluded h = 5 images both before and after the test set to account for autocorrelation; this value was chosen based on observations that the strongest autocorrelation occurs within 5 images (with standard TR values of 2 s as we have here), with little influence outside of that window (Bullmore et al., 1996). Each test set consisted of v = 11 images (one image per TR). The training set included g = 40 images before and after the hold-out data chosen arbitrarily (although we do not expect results to depend strongly on this choice, it could be optimized in future, larger studies). These combinations of 40 + 5 + 11 + 5 + 40 images were “rolled” across the dataset of 266 images without overlap of the test sets, leading to 24 cross-validation folds. This resulted in using only 264 of the 266 TRs. To keep the size of the training and the test set equal across folds and to be able to include the first and last 45 TRs in test sets, we moved the test set along the data starting at the first TR. Therefore, the first fold included, in order, 11 TRs for the test set (v), 5 TRs for the excluded images (h), and 80 (40 + 40) TRs for the training set (g). The last fold was 5 TRs for the training set, followed by five TRs for the excluded images and 11 TRs for the test set. With this cross-validation scheme, every image is used in the test set exactly once and training and testing set sizes are equal across all folds.

**Group based cross-validation.** To assess how well the average group map could predict a new participant’s ANS signals, we used a leave-one-subject-out cross-validation procedure. For each test subject, we calculated a predicted physiological time series by applying the following steps. First, in the training phase, we estimated each training participant’s individual weight map (not including the test individual). Then, we averaged the training weight maps across training participants into a single predictive map. Third, in the test phase, we applied the predictive map to each volume of the test subject’s fMRI time series, calculating the spatial correlation between the predictive map and each test image to generate a predicted physiological time series. After estimating the test participant’s predicted physiological response, we calculated the correlation between the predicted and actual physiological time series. We repeated this procedure for each test (held-out) participant and report the average correlation between predicted and actual time courses across all test subjects in the cross-validation procedure (Fig. 2B).
parable to the HRF (Berntson et al., 1997). Although parasympathetic contributions have a short time constant and can affect the lag of the next heartbeat (e.g., within 1 s), sympathetic influences generally occupy a lower-frequency range within 0.05–0.15 Hz (Berntson et al., 1993; Berntson et al., 1997; Montano et al., 2001), so sympathetic outflow has a time constant of ~4–6 s (Eckberg et al., 1985). There are some potential differences in relative lag of the fMRI–HR relationship across regions (Wager et al., 2009a), but these are small. Here, we tested the influence of potential lags by retraining our multivariate models using varying lags shifted ±1 to ±35 TRs.

Thresholding. The pattern maps encompass voxels across the brain. To find the weights that contributed most reliably to the overall HR/SCL prediction across participants, we performed a univariate one-sample t test on the pattern weights for each voxel. This allowed us to test whether each voxel’s contribution to predicting ANS activity was significantly different from zero across participants. We present the results of these statistical maps thresholded using p < 0.001 uncorrected for display. The mPFC was specifically of interest, given both previous human imaging work on correlates of autonomic responses (Gianaros and Sheu, 2009; Beissner et al., 2013) and recent work showing anatomical projections from multiple medial PFC zones to the peripheral sympathetic nervous system in monkeys (Dum et al., 2016). To control the familywise error rate, we additionally performed a nonparametric permutation analysis using 5000 samples (Nichols and Holmes, 2002; Maris and Oostenveld, 2007) with threshold-free cluster enhancement (Smith and Nichols, 2009) in areas of the mPFC that were specifically of interest to us. We used the anterior and middle mPFC masks from a recent three cluster meta-analytic parcellation of the mPFC, which contained 6010 and 3054 voxels, respectively (de la Vega et al., 2016). The ROI masks were selected a priori and are publically available on neurovault (http://www.neurovault.org/collections/1458/, RRID:SCR_003806). Cross-prediction analyses focus on the predictions made by the entire cross-validated pattern and are thus not subject to multiple-comparisons concerns because only one test is performed. As an additional way to localize analyses without requiring voxelwise multiple comparisons, we also calculated spatial similarity with known resting-state networks, described below.

Additional a priori regions of interest (ROIs). In addition to the predictive analyses, we focused inferences about brain similarity on specific ROIs based on previous work on autonomic generators in animal models (Cechetto and Saper, 1990; Yasui et al., 1991; Saper, 2002; Price and Dreves, 2010) and human work (Critchley, 2002; Gianaros et al., 2005; Wager et al., 2009a; Beissner et al., 2013; Wager and Gianaros, 2014; Dum et al., 2016) focusing on the medial prefrontal–brainstem axis. The most reliable correlates of autonomic activity detectable using fMRI at this resolution include increased activity (positive brain-autonomic correlations) in dorsal anterior cingulate cortex (dACC) and periaqueductal gray (PAG) and reduced activity in ventromedial PFC (vMPFC). We used the anterior division of the cingulate gyrus dilated 5 mm from the Harvard-Oxford Probabilistic Cortical Structural Atlas (Desikan et al., 2006) and a manually drawn mask of the vMPFC.

Cross-modality prediction analyses. Validating autonomic responses on other types of autonomic activity (e.g., correlating SCL with HR) assumes that each signal results from a single underlying factor (stress or arousal) plus noise. Validation using an external measure such as fMRI can help determine whether responses to stressors or emotions result from common or separable factors.

If cross-modality prediction (e.g., a model trained on HR tested on SCL) is as strong as within-modality prediction (e.g., trained and tested on HR), then there is evidence for a single common factor underlying both measures. For example, if HR and SCL are equally good indicators of “stress” and cerebral activity predictive of HR is caused by activation of a unitary stress response, then both signals should be predicted equally well by either an HR- or an SCL-related brain pattern. Within-modality prediction can thus establish an upper bound on the amount of signal that can be explained relative to the noise and provides an expected value for the amount of variance that should be explained if both physiological measures are equally good indicators of a single underlying construct.

Alternatively, if within-modality prediction is substantially stronger than cross-modality prediction, that difference constitutes evidence against the unitary factor model and suggests that each signal is measuring a separable factor. For example, if each autonomic signal can only be predicted by their respective brain patterns and there is no evidence of cross-prediction, then each pattern is separately modifiable and indicates no degree of common variance between the two factors (Bechara et al., 1995; Plaut, 1995; Sternberg, 2001). Therefore, cross-modality prediction analyses can allow us to make
inferences about whether HR and SCL measure a single underlying response to stress or distinct aspects of the stress response.

HR versus SCL weight map contrast. To investigate differences between the pattern maps for HR and SCL independently of other brain areas, we computed a difference map by subtracting the SCL-related map from the HR-related map. This statistical difference map was then thresholded using \( p < 0.001 \) uncorrected. We focused on dACC, vmPFC, and PAG.

Local spatial similarity analyses. To investigate whether each pattern predictive of autonomic signals contained overlapping brain signals, we conducted a local spatial similarity analysis. This entailed calculating the spatial correlation between the two patterns for each participant using a spherical searchlight with a 5-voxel radius. We then performed a one-sample \( t \) test to determine which searchlight correlations were reliably significantly different from zero across participants. The resulting statistical map was thresholded using \( p < 0.001 \) uncorrected, focusing on a priori ROIs. Therefore, inferences in regions not predicted based on prior work are preliminary and we focus primarily on our strongly predicted ROIs.

Large-scale spatial pattern similarity analysis. In addition, we compared the weight maps with seven known networks derived from a large sample (\( N = 1000; \) Buckner et al., 2011) using spatial pattern similarity analysis. We then computed point–biserial correlations between the networks and the HR/SCL weights maps and \( t \) tests on Fisher’s transformations of those correlation coefficients.

Results

Autonomic physiology

As reported previously (Wager et al., 2009a; Wager et al., 2009b), the social threat task increased ANS measures reliably compared with baseline. HR increased during speech preparation (\( post \) hoc comparison for HR before vs during speech task: \( t_{(17)} = -6.04, p < 0.001 \) ) and returned back to baseline after the stress (\( post \) hoc comparison for HR during vs after speech task: \( t_{(17)} = 0.74, p = 0.47 \) ). SCL increased nonsignificantly during speech (\( post \) hoc comparison for SCL before vs during speech task: \( t_{(17)} = -0.94, p < 0.36 \) ) and decreased significantly after the stress (\( post \) hoc comparison for SCL during vs after speech task: \( t_{(17)} = 3.49, p = 0.003 \) ). HR and SCL were modestly but significantly correlated across the entire time course (\( r = 0.23 \pm 0.09 \) (SEM), \( t_{(17)} = 2.45, p = 0.03 \); Figure 3).

Cross-validated prediction: accuracy

We first present results on the accuracy of whole-brain within-modality prediction and cross prediction. Subsequently, we present neuroscientific results on the predictive maps and their similarity in ROIs.

Within-modality prediction

The time course of the true HR and SCL and the cross-validated predicted time courses based on idiographic (within-person) training are shown in Figure 3. For HR, the mean correlation between predicted and true values across time was \( r = 0.54 \pm 0.05 \) (\( t_{(17)} = 12.06, p < 0.001 \) ). For SCL, the mean correlation between predicted and true values across time was \( r = 0.58 \pm 0.06 \) (\( t_{(17)} = 9.45, p < 0.001 \); Fig. 3). Therefore, within-modality prediction of each measure based on whole-brain patterns of brain activity was relatively strong (see time course panels on the right of Fig. 3), establishing that brain activity patterns are sufficient to predict HR and SCL levels on out-of-sample data. They are also comparably strong, indicating that neither measure was substantially more strongly related to brain activity.

The predictive maps for HR and SCL were modestly but significantly spatially correlated (\( r = 0.29 \pm 0.12, t_{(17)} = 2.40, p = 0.03 \) ) across subjects, comparable to the correlation between HR and SCL.

Cross-prediction

Cross-prediction of each measure using the predictive map trained on the other measure (i.e., training on HR and testing on SCL, and vice versa) can help to establish how much of the within-modality relationship is attributable to a common, nonmeasure-specific factor. Cross-prediction was significant for both measures: \( r = 0.24 \) (\( t_{(17)} = 2.42, p = 0.03 \) ) for training on HR and testing on SCL and \( r = 0.28 \) (\( t_{(17)} = 2.91, p = 0.01 \) ) for training on SCL and testing on HR. However, they were substantially weaker than within-modality prediction (HR: \( t_{(17)} = 2.61, p = 0.02 \); SCL: \( t_{(17)} = 2.95, p = 0.01 \), indicating that the predictive brain maps are capturing appreciable measure-specific variance.

Figure 3. Correlations and SDs between observed and predicted HR and SCL and cross-prediction results with observed (black) and idiographically predicted HR (green) and SCL (blue) (z-scored).
Temporal considerations
To investigate the influence of temporal lag on the prediction, we ran cross-validated analyses with temporally shifted data. For a shift of SCL toward later stages of the BOLD response, correlations between predicted and observed SCL significantly drop after shifting more than 20 TRs (Fig. 5, top left) and, for a shift of HR toward later stages of the BOLD response, correlations between predicted and observed HR were not significant (Fig. 5, top right). These findings suggest that predictive patterns do not capture baseline physiological fluctuations in the absence of stress in at least some cases. However, they do not depend exclusively on active thought processes occurring during speech preparation because they are also present during the poststress epoch. Overall, the relationship between the brain patterns observed and physiological measures appears to be stress related and not strongly dependent on active speech preparation.
toward earlier stages of the BOLD, correlations significantly drop after 25 TRs. Similarly, for a shift of HR toward later stages of the BOLD response, correlations drop after 6 TRs (Fig. 5, bottom left) and, for a shift toward earlier stages of the BOLD response, correlations only drop significantly at more than 35 TRs. These results show that the brain-autonomic correlations were fairly insensitive to the relative timing of physiology and BOLD in this dataset, which might be expected given the low temporal frequency nature of the task.

Comparison with group-based prediction
The previous analyses have focused on investigating the predictive value of idiographic (within-participant) weight maps within and across measures. An additional question is to assess the stability of these weight maps across participants by evaluating how well the group weight map generalizes to new participants. Using leave-one-subject-out cross-validation, we found that both group maps significantly predicted the time course of ANS responses in new participants, but substantially weaker than idiographic prediction. HR-related brain patterns predicted both HR \(r = 0.32 \pm 0.07, t_{17} = 4.57, p < 0.001\) and SCL \(r = 0.24 \pm 0.07, t_{17} = 3.71, p = 0.002\) equally strongly, showing comparable accuracy for within-measure prediction and cross-prediction. Likewise, SCL-related brain patterns predicted both SCL \(r = 0.28 \pm 0.08, t_{17} = 3.67, p = 0.002\) and HR \(r = 0.27 \pm 0.07, t_{17} = 3.63, p = 0.002\) equally strongly (Fig. 4B). Therefore, individualized predictive maps explain more variance in physiology than group maps and are required to obtain autonomic measure-specific brain maps. This suggests that group-level pre-

---

**Figure 5.** Correlations (and their SDs) between predicted and observed HR and SCL based on shifted data. Negative shifts \((-1 \text{ to } -35)\) represent shifts in which HR and SCL data were aligned to later TRs of the BOLD response, whereas positive shifts \((+1 \text{ to } +35)\) represent shifts in which HR and SCL data were aligned to earlier TRs of the BOLD response. Black indicates zero-shift correlation related to the use of temporally aligned HR/SCL and BOLD data, which was used for the main analyses. Correlations (and their SDs) between predicted and observed SCL (top) and HR (bottom) are based on shifted data. Negative shifts \((-1 \text{ to } -35)\) represent shifts in which HR and SCL data were aligned to later TRs of the BOLD response, whereas positive shifts \((+1 \text{ to } +35)\) represent shifts in which HR and SCL data were aligned to earlier TRs of the BOLD response. Black indicates zero-shift correlation related to the use of temporally aligned HR/SCL and BOLD data, which was used for the main analyses. Asterisks mark significant t tests comparing the zero-shift correlation coefficient with the shifted correlation coefficients. ***\(p < 0.001\), **\(p < 0.01\), *\(p < 0.05\).
diction is less measure specific and likely reflects central mechanisms of threat appraisal related to both measures.

Cross-validated prediction: characteristics of predictive patterns

Pattern characteristics and ROIs
Although the entire pattern was used for prediction, weights in several areas contributed consistently to the prediction of autonomic responses across participants (Fig. 6). HR was predicted by positive weights in the dorsal and posterior parts of the cingulate and cerebellum and negative weights in the vPFC and dmPFC hippocampus and temporal pole (Fig. 6, Table 1). SCL, in contrast, was predicted by positive weights in the rostromedial PFC and cerebellum and negative weights in the vPFC and dmPFC (Fig. 6, Table 2). We conducted additional analyses to identify regions that were both unique and common to the prediction of each autonomic signal.

Specificity of the predictive pattern for HR versus SCL
To identify regions that uniquely contributed to the prediction of HR and SCL across participants, we conducted univariate paired $t$ tests between the within-subject HR and SCL maps. The analyses above identify some commonalities in brain patterns predictive of HR and SCL, indicative of shared representation. Next, we sought to test for significant differences, which shed light on the differential representation of HR and SCL in the brain. We estimated an HR versus SCL difference map by subtracting the SCL pattern map from the HR pattern map and repeating this procedure with bootstrap resampling to obtain $p$-values and make in-

![Figure 6. Predictive weight maps for HR and SCL. Maps are thresholded at $p < 0.001$ (one-tailed) for display and interpretation, but the entire whole-brain patterns were used in prediction. Warm colors (orange/yellow) indicate increased predicted physiology with increasing brain activity and cool colors (blue) indicate decreased predicted physiology with increasing brain activity. Top, Medial surface maps for HR (left) and SCL (middle) are shown, with the full patterns in dACC and vmPFC (two midline structures of a priori interest), as well as the difference map for HR − SCL (right). Green outlines show the boundaries of Familywise Error Rate corrected results ($p < 0.05$) based on a permutation test. Bottom, Montages of full maps showing axial slices from −40 to 50 mm in 6 mm intervals for HR and SCL and for the difference HR − SCL. Positive predictive weights were found for both measures in dACC and superior cerebellum and negative weights were found in vmPFC and temporal pole. However, the patterns were not strongly correlated in many other regions, indicating significant divergence between predictive maps for the two patterns. Specifically, vmPFC and insula decreases and lateral prefrontal and pre-SMA increases were more strongly predictive of HR than SCL, whereas dmPFC increases were more strongly predictive of SCL (axial slices at the bottom).]

**Table 1. HR-related weight map peak regions with positive and negative effects**

<table>
<thead>
<tr>
<th>Region</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>Volume (mm$^3$)</th>
<th>$t$ (df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal gyrus</td>
<td>−50</td>
<td>−14</td>
<td>−36</td>
<td>34</td>
<td>272</td>
<td>7.55 (17)</td>
</tr>
<tr>
<td>vmPFC</td>
<td>10</td>
<td>42</td>
<td>−30</td>
<td>1</td>
<td>8</td>
<td>−5.18 (17)</td>
</tr>
<tr>
<td>vmPFC</td>
<td>14</td>
<td>50</td>
<td>−24</td>
<td>2</td>
<td>16</td>
<td>−5.15 (17)</td>
</tr>
<tr>
<td>Lateral OFC</td>
<td>−20</td>
<td>40</td>
<td>−28</td>
<td>1</td>
<td>8</td>
<td>−5.84 (17)</td>
</tr>
<tr>
<td>OFC</td>
<td>14</td>
<td>44</td>
<td>−24</td>
<td>16</td>
<td>128</td>
<td>−6.21 (17)</td>
</tr>
<tr>
<td>OFC</td>
<td>−20</td>
<td>38</td>
<td>−20</td>
<td>88</td>
<td>704</td>
<td>−7.15 (17)</td>
</tr>
<tr>
<td>Lateral OFC</td>
<td>−24</td>
<td>42</td>
<td>−26</td>
<td>1</td>
<td>8</td>
<td>−5.11 (17)</td>
</tr>
<tr>
<td>Medial OFC</td>
<td>10</td>
<td>50</td>
<td>−12</td>
<td>2</td>
<td>16</td>
<td>−5.31 (17)</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>0</td>
<td>48</td>
<td>2</td>
<td>16</td>
<td>808</td>
<td>−5.57 (17)</td>
</tr>
<tr>
<td>dIPFC</td>
<td>46</td>
<td>16</td>
<td>48</td>
<td>23</td>
<td>184</td>
<td>6.27 (17)</td>
</tr>
<tr>
<td>ACC</td>
<td>8</td>
<td>38</td>
<td>−10</td>
<td>101</td>
<td>808</td>
<td>−6.76 (17)*</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>2</td>
<td>−52</td>
<td>−24</td>
<td>1</td>
<td>8</td>
<td>5.17 (17)</td>
</tr>
<tr>
<td>Pons</td>
<td>−12</td>
<td>−22</td>
<td>−38</td>
<td>3</td>
<td>24</td>
<td>−5.76 (17)</td>
</tr>
</tbody>
</table>

$x$, $y$, $z$, MNI coordinates; $t$, peak $t$-value, thresholded at $0.001$. *Surviving FWE correction.
ferences about which brain regions significantly contributed to differentially predicting physiology.

The results of this analysis highlight several regions that were more strongly predictive of HR than SCL at \( p < 0.001 \) uncorrected. Weights in the dACC and lateral prefrontal PFC were more positively associated with HR, whereas weights in the subgenual ACC and dmPFC were more negatively associated with HR relative to SCL. These results indicate that HR and SCL are to some degree represented differentially across the brain, particularly in the dACC, dmPFC, and vmPFC.

In addition to the whole-brain analyses, we also compared local predictive patterns within ACC and vmPFC ROIs. Local pattern responses reflect the local contribution of an ROI to the overall prediction, controlling to some extent for activity in other brain regions. The ACC parts of the weight maps for HR and SCL were not significantly correlated (\( r = 0.04, t_{17} = 0.68, p = 0.51 \)), whereas local vmPFC patterns showed strong correlations across the two maps (\( r = 0.74, t_{17} = 3.50, p = 0.003 \)). This pattern of findings indicates potential shared representations underlying both measures in vmPFC and distinct representations in ACC.

Commonality of the predictive pattern for HR versus SCL: local spatial similarity analyses

Although the HR and SCL patterns appear to have some degree of differential representation across the brain, there are likely other regions that have a more shared representation. To test local pattern similarity throughout the brain, we used a “searchlight” (spatial moving window) with a 5-voxel radius to identify brain regions that had consistently strong correlations between HR and SCL predictive patterns across participants. The goal of this analysis was to identify regions that had common shared representations across the two patterns.

Results show the strongest correlated local patterns in vmPFC extending into lateral OFC, anterior temporal cortices, parahippocampal cortex, medial thalamus, paracentral lobule extending into precuneus, temporoparietal junction, and brainstem extending into the approximate area of the locus ceruleus and PAG (Fig. 7, Table 3). These analyses identify common representations in a number of limbic and paralimbic cortical regions associated with emotion, stress, and social cognition in previous studies (Gianaros and Sheu, 2009; Frith and Frith, 2012; Lindquist et al., 2012).

Pattern overlap with known networks: large-scale spatial pattern similarity analysis

To test whether this pattern of potential common representations is associated with specific previously identified cortical networks, we calculated the similarity of the individual searchlight maps with seven resting-state networks identified in a large sample of 1000 individuals (Buckner et al., 2011; Yeo et al., 2011; Choi et al., 2012). Figure 7 shows the mean correlation with SEs across individual participants for each of the seven networks. As shown in Figure 7, the regions with correlated predictive patterns mapped almost exclusively onto the “limbic” network identified in the resting-state studies (Yeo et al., 2011), with a significant spatial correlation with the limbic network (\( r = 0.03, p = 0.01 \)) and nonsignificant correlations with other networks, except visual (\( r = -0.02, p = 0.01 \)).

Discussion

This study used machine learning to investigate the brain bases of HR and SCL responses during social threat. Two brain patterns were strongly predictive of the time course of HR and SCL. Each pattern weakly predicted the other measure, indicating some common brain representation and a substantial divergence in the brain bases of each autonomic measure. Analysis of local pattern similarity revealed correlated patterns, implying shared representation, in a “limbic” network including vmPFC and temporal and parahippocampal cortices and portions of the thalamus and brainstem (covering PAG and locus ceruleus) known to be anatomically interconnected with this limbic cortical network and to drive autonomic responses. Together, these results provide a novel picture of shared and measure-specific brain representation of autonomic responses during stress, which is relevant for understanding the unique health risks associated with specific types of autonomic reactivity (Patel et al., 1995; Campbell and Ehlert, 2012).

One system or many?: implications for patterned autonomic responses

A pervasive hypothesis over the past century is that the ANS operates as a unitary system when responding to stressors (Cannon, 1935; Lazarus et al., 1963; Grandjean et al., 2008; Valenza et al., 2012). However, there is also substantial evidence indicating that outputs to different body systems (e.g., heart, skin, pupils, bone marrow, and other organs) operate relatively independently (Jä nig and McLachlan, 1992; Norman et al., 2014). For example, \( \beta \) antagonists block increases in HR, but not SCL, during stress (Jacobs et al., 1994).

One question that has been difficult to address in human studies is whether unshared variance across HR and SCL (e.g., 86% unshared in Croft et al., 2004) reflects measurement noise or systematic differences indicative of true patterned responses (Levenson, 2014). We observed similarly low correlations between HR and SCL, but also that each measure was much more reliably predicted by an external referent, in this case brain activity. Therefore, the unshared variance across HR and SCL is not noise; a substantial amount reflects true patterned autonomic responses with an identifiable brain basis.

Among other differences, we found that predictive patterns in the dACC are uncorrelated across measures. Furthermore, dACC is more predictive of HR, whereas the vmPFC is more predictive of SCL (Fig. 4). Our results are incompatible with a strong version of unitary theories of stress-induced autonomic responses and suggest that there is differential involvement of sympathetic and parasympathetic divisions in HR and SCL (Berntson et al., 1994). Sympathetic measures such as SCL likely have a similar delay as the HRF, whereas parasympathetic contributions to HR are thought to be more rapid (Berntson, Cacioppo, and Quigley, 1993; Berntson et al., 1997; Montano, Porta, and Malliani, 2001). Therefore, it is possible that the observed differences between SCL- and HR-related patterns could reflect differential involvement of parasympathetic and sympathetic responses (Berntson et al., 1997).

Multiple brain systems

The brain patterns that predicted autonomic activity were distributed across the brain, including regions in brainstem, me-

---

**Table 2. SCL-related weight map peak regions with positive and negative effects**

<table>
<thead>
<tr>
<th>Region</th>
<th>( x )</th>
<th>( y )</th>
<th>( z )</th>
<th>Voxels</th>
<th>Volume (mm(^3))</th>
<th>( t ) (df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateral OFG</td>
<td>−26</td>
<td>24</td>
<td>−30</td>
<td>4</td>
<td>32</td>
<td>−7.55 (17)*</td>
</tr>
<tr>
<td>Lateral OFG</td>
<td>−22</td>
<td>24</td>
<td>−26</td>
<td>7</td>
<td>56</td>
<td>−6.21 (17)</td>
</tr>
<tr>
<td>dACC</td>
<td>−20</td>
<td>8</td>
<td>52</td>
<td>8</td>
<td>64</td>
<td>5.84 (17)</td>
</tr>
</tbody>
</table>

\( x, y, z \), MNI coordinates; \( t \), peak \( t \)-value; thresholded at 0.001.

*Surviving FWE correction.
dorsal and lateral prefrontal, temporal, and parietal cortices. This extends an autonomic regulation network of brainstem and some subcortical regions such as the hypothalamus (Saper, 2002; Price et al., 2008) to areas of the PFC thought to be involved in autonomic regulation (e.g., vmPFC, dACC, and insula; see Critchley and Harrison, 2013) and some with a role in autonemics that is less well understood (e.g., lateral PFC). Interestingly, recent neuroanatomical work in monkeys has shown multisynaptic projections to a peripheral sympathetic effector system (the adrenal medulla) from multiple discrete areas within the medial wall, including dorsal and ventral aspects of the mPFC and SMA (Dum et al., 2016). This diversity of regions parallels recent findings that emotional states are not only related to activity in single areas of the brain, but also to a complex network (Kassam et al., 2013; Nummenmaa et al., 2014; Kragel and LaBar, 2015).

Our HR-predictive patterns, in particular, also identified multiple areas of the dorsal and ventral mPFC and SMA (Fig. 6), as in the recent study by Dum et al. (2016). Interestingly, the weight maps for HR and SCL also differ in another respect: the pattern for HR includes larger clusters compared with the pattern for SCL. Because we also show that SCL can be predicted as well as HR, we cannot assume that these differences are based on methodological issues, but the HR-related pattern seems to be more localized, capturing more variance, whereas the pattern for SCL is more distributed across systems.

Further, the comparison analyses with established networks (Yeo et al., 2011) and the spatial similarity maps of HR and SCL predictive maps showed an overlap with the resting-state-derived network for limbic activity that is relevant for emotional responding such as in the context of acute stress, validating the emotional stress-related nature of the predictive maps.

Figure 7. Spatial similarity analysis showing areas with significant local pattern correlations for HR- and SCL-predictive patterns. These results complement the previous analyses showing the strongest correlated local patterns in vmPFC extending into lateral OFC, anterior temporal cortices, medial thalamus, paracentral lobe extending into precuneus, temporoparietal junction, brainstem extending into area covering LC and PAG and parahippocampal cortex (left). Comparing the spatial similarities with known networks (Buckner et al., 2011; Yeo et al., 2011) shows a significant overlap with the limbic network.

Table 3. Peak regions with positive and negative effects for representational similarity analysis

<table>
<thead>
<tr>
<th>Region</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>Volume (mm³)</th>
<th>t (df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle temporal gyrus</td>
<td>56</td>
<td>8</td>
<td>−18</td>
<td>17</td>
<td>136</td>
<td>4.53 (17)</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>46</td>
<td>−44</td>
<td>18</td>
<td>49</td>
<td>392</td>
<td>6.06 (17)</td>
</tr>
<tr>
<td>Inferior parietal lobe</td>
<td>40</td>
<td>−36</td>
<td>44</td>
<td>37</td>
<td>296</td>
<td>5.66 (17)</td>
</tr>
<tr>
<td>Paracentral lobule</td>
<td>−2</td>
<td>−38</td>
<td>64</td>
<td>16</td>
<td>128</td>
<td>4.6 (17)</td>
</tr>
<tr>
<td>Medial orbitofrontal cortex</td>
<td>−4</td>
<td>42</td>
<td>−30</td>
<td>19</td>
<td>152</td>
<td>5.18 (17)</td>
</tr>
<tr>
<td>Insula</td>
<td>46</td>
<td>0</td>
<td>−26</td>
<td>13</td>
<td>104</td>
<td>5.32 (17)</td>
</tr>
<tr>
<td>Anterior insular cortex</td>
<td>−56</td>
<td>18</td>
<td>−16</td>
<td>21</td>
<td>168</td>
<td>4.39 (17)</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>12</td>
<td>−54</td>
<td>−44</td>
<td>11</td>
<td>88</td>
<td>4.96 (17)</td>
</tr>
</tbody>
</table>

x, y, z, MNI coordinates; t, peak t-value, thresholded at 0.001.

Context dependence and limits on generalizability

The current results are based on a single social threat paradigm that has been widely studied because it can create powerful physiological responses in a laboratory environment and is ethically possible (Dickerson and Kemeny, 2004). These predictive patterns’ generalizability to other emotional situations must be tested empirically in future studies. The patterns that we identified are likely to be related to stress, but not driven solely by the cognitive demands of speech preparation per se because they still correlate with HR and SCL in the post-task phase, when the demand to prepare the speech is removed but many individuals feel residual stress. In addition, the brain areas most closely related to autonomic responses in this study—including vmPFC, dACC, insula, amygdala, and others—are related to autonomic changes across diverse challenges in human studies (Gianaros and Sherr, 2009; Beissner et al., 2013; Chang et al., 2015; Gianaros and Wagner, 2015) and animal models of stress (Saper, 2002).

To more fully test the context dependency of brain–autonomic relationships, studies involving different types of tasks are needed, particularly measuring multiple types of autonomic activity moment by moment, as we did here. This will also contribute to clarifying the effect of time lags. We observed some invariance in the predictive validity within a range of ~80 s. We believe this lack of specificity at smaller temporal shifts can be attributed to the slow-frequency design of the study (<.01 Hz). Future work should further investigate this question with a design more amenable to identifying the specificity of the temporal response.

Limitations

A benefit of the current approach is that it tracks within-person, moment-by-moment variation in autonomic responses across time (Levenson, 2014). This within-person approach is more robust to systematic noise caused by heterogeneity across participants in brain size and shape and nuisance factors affecting BOLD signal magnitude (e.g., hematocrit levels and caffeine intake) than individual differences approaches. However, our study is limited to one type of stress and the generalizability to other stressors and emotion-related remains unknown.

Another endemic issue when studying physiology is the possibility that HR-related activity reflects physiological (e.g., vascular, motion) artifacts rather than neural/glial changes of interest. Although this is possible to some extent, we think that such confounds are unlikely to drive our results for several reasons. First, the regions that we identified as being the most important for prediction are known to be critical for neural control of autonomic in animal models and human stimulation and lesion.
studies (Saper, 2002; Critchley et al., 2003). Second, vascular confounds typically involve patterns that include the borders of fluid spaces and large vessels, particularly around the base of the brain, which we did not observe here (Glover et al., 2004). Third, different patterns of brain activity are differentially predictive of HR and SCL levels, which is incompatible with global confounds.

In addition, we did not separate cardiac effects related to versus unrelated to the respiratory cycle. HR is linked physiologically to respiration and these relationships seems to be mediated by blood CO2-sensing systems in the brain (Porges and Byrne, 1992; Cacioppo et al., 1994; Brooks et al., 2013) and likely in part through the baroreceptor reflex circuit (Gianatos and Wager, 2015). These respiration effects on HR can sometimes be artifacts (i.e., related to chest wall motion), but are widely thought to be largely nonartifactual, reflecting complex interactions between the respiratory and cardiovascular systems. Therefore, we opted not to try to separate out respiratory effects here.

Conclusion

In summary, we found evidence for both common and measure-specific brain patterns that predict physiological responses to social threat. These findings suggest that social threat does not result in a unitary autonomic response, but rather in multiple autonomic responses in different body systems that are largely decoupled, but nonetheless related to identifiable patterns of brain activity. These findings are relevant not only for understanding stress responses and how they should be investigated in future research, but also for our understanding of how coherent physiological measures of autonomic activity really are. They are also relevant for identifying brain patterns related to specific autonomic measures that confer specific types of health risks.

References


