

On the Role of Situational Stressors in the Disruption of Global Neural Network Stability during Problem Solving

Mengting Liu, Rachel C. Amey, and Chad E. Forbes

Abstract

■ When individuals are placed in stressful situations, they are likely to exhibit deficits in cognitive capacity over and above situational demands. Despite this, individuals may still persevere and ultimately succeed in these situations. Little is known, however, about neural network properties that instantiate success or failure in both neutral and stressful situations, particularly with respect to regions integral for problem-solving processes that are necessary for optimal performance on more complex tasks. In this study, we outline how hidden Markov modeling based on multivoxel pattern analysis can be used to quantify unique brain states underlying complex network interactions that yield either successful or unsuccessful problem

solving in more neutral or stressful situations. We provide evidence that brain network stability and states underlying synchronous interactions in regions integral for problem-solving processes are key predictors of whether individuals succeed or fail in stressful situations. Findings also suggested that individuals utilize discriminate neural patterns in successfully solving problems in stressful or neutral situations. Findings overall highlight how hidden Markov modeling can provide myriad possibilities for quantifying and better understanding the role of global network interactions in the problem-solving process and how the said interactions predict success or failure in different contexts. ■

INTRODUCTION

When individuals are placed in stressful situations, they tend to exhibit performance deficits (Beilock, 2008). Performance decrements likely stem from the interaction between a variety of cognitive processes; however, little is known about how global neural network stability that instantiates different cognitive processes is modulated under stress or what properties of network stability are associated with both successful and unsuccessful performance. The purpose of this study is to utilize multivoxel pattern analysis (MVPA) and hidden Markov modeling (HMM) to provide a novel means of identifying and quantifying unique changes in global neural network activity that underlie successful and unsuccessful problem solving in stressful compared with neutral environments. We provide evidence that brain network stability and synchronous activity in regions integral for problem solving are key predictors of whether individuals succeed or fail in stressful situations. Findings inform our understanding of how global neural network stability is altered when stressed or nonstressed individuals complete cognitively demanding tasks as well as what properties of network stability may ultimately promote success in stressful situations.

Solving cognitive tasks in stressful situations recruits a multitude of physiological and cognitive processes that ultimately engender underperformance, including in-

creased cortisol (Takahashi et al., 2005), taxation of working memory resources, increased emotion regulation (Beilock, 2008; Schmader, Johns, & Forbes, 2008; Ashcraft & Kirk, 2001; Ashcraft & Faust, 1994), and reduced attentional control (Suarez-Pellicioni, Nunez-Pena, & Colome, 2013, 2014; Beilock & Gray, 2012; Eysenck, Derakshan, Santos, & Calvo, 2007; Hopko, McNeil, Gleason, & Rabalais, 2002). At the neural level, this cascade of simultaneous physiological and cognitive responses manifests as altered activity in brain networks integral for working memory, emotion, and attentional processes and thus problem solving in general. Problem solving itself requires coordination between neural networks that instantiate varied cognitive processes. For example, past research suggests that connectivity between frontoparietal brain regions increases during arithmetic tasks (Rosenberg-Lee, Barth, & Menon, 2011), intelligence test tasks (Vakhtin, Ryman, Flores, & Jung, 2014), and visual-spatial working memory tasks (Sauseng, Klimesch, Schabus, & Doppelmayr, 2005). Solving problems in stressful contexts appears to alter connectivity between regions integral for attentional and emotional processes. For example, when individuals were placed in socially threatening situations, they exhibited decreased connectivity between the inferior parietal sulcus and dorsolateral pFC (DLPFC) during a working memory task (the *n*-back task) and increased connectivity between the inferior parietal sulcus and the ventromedial pFC, a region often implicated in the regulation of autonomic and emotional responses (van Ast et al., 2016). Alterations in connectivity within a frontoparietal network that

mediates attention shifts and the amygdala have also been found (Henckens, van Wingen, Joels, & Fernandez, 2012; Liston, McEwen, & Casey, 2009). These findings suggest that solving problems in stressful contexts likely alters functional connectivity between neural networks integral for multiple cognitive processes and thus perhaps globally across the brain.

Recent theoretical and empirical studies have stressed the importance of understanding neural function and cognition from a global network perspective (Sporns, 2013; Liu, Kuo, & Chiu, 2011). One way that this may be possible is via measuring network stability (Rzucidlo, Roseman, Laurienti, & Dagenbach, 2013; Telesford et al., 2010), where a network is considered stable if similar patterns are repeatedly found when a given brain network is active during a given task. To measure this, it would be necessary to capture the dynamic and flexible nature of neural activity within and between regions across the brain; however, current approaches typically reflect a static and time-invariant approximation of underlying neural dynamics (Bola & Borchardt, 2016). It is highly likely that, over the course of a cognitively demanding task, the brain exhibits myriad transient states, oscillating in and out of synchrony (and thus communication) within and between multiple brain regions at both slow and fast frequencies and possibly consisting of functional networks that transiently shift toward more globally integrated arrangements (Bola & Sabel, 2015; Buzsáki, 2006). Thus, it would be necessary to identify an analytic strategy that captures both global network stability and transient brain states on the order of milliseconds across the full time scale of a given cognitive process.

MVPA and HMM provide a novel means to address these issues. MVPA attempts to map distributed patterns of brain activity onto various psychological categories (Norman, Polyn, Detre, & Haxby, 2006). In typical MVPA, brain patterns are segregated and classified into events with fixed time intervals. However, during a given cognitive process like problem solving, brain patterns are likely spontaneously and dynamically changing on the order of milliseconds over the course of seconds; thus, one has to determine how to divide the temporal continuous data into discrete epochs to be classified, which likely collapses across multiple neural events important for the problem-solving process (Anderson, Fincham, Schneider, & Yang, 2012). Here, HMM can be utilized to determine exactly how to parse apart a given cognitive process into different temporal states. HMM assumes that any real-world process is composed of one or more temporally distinct intrinsic stages or “hidden states” (Chiu et al., 2011; Rabiner, 1989). With respect to neural activity, HMM can simultaneously segment and classify hidden multivoxel patterns/states that represent unique, meaningful changes in neural activity within and between neural regions that occur during execution of a given cognitive process. Once the HMM–MVPA model is established, it is possible to obtain such measures as how long

the brain spends in a specific state, how often the brain reenters a specific neural state, and how many states appear in a given temporal range overall. All of these indices can be used to gain insight into how dynamic or stable different neural regions are in relation to one another. If a given process was stable across time, for example, activity within brain regions does not change over the course of a single trial in relation to one another, HMM would categorize the entire process as one single state. Conversely, if the degree of stability in a given process is low, for example, activity within brain regions changes frequently over the course of a single trial in relation to one another, HMM would categorize the entire process as consisting of several unique states. HMM thus provides a means to quantify global network activity and stability that otherwise would be impossible to detect in raw signal or via other functional connectivity analytic approaches (e.g., psychophysiological interaction analyses). Initial attempts at using HMM on brain data have revealed that problem solving can be decomposed into several states consisting of unique neural events (Borst & Anderson, 2015; Anderson, Lee, & Fincham, 2014; Anderson et al., 2012). Whether these states map onto global network stability or predict specific behaviors in different contexts are questions unexplored in the literature to date.

In this study, we utilized HMM to dynamically model global neural network stability while individuals solved math problems ranging in difficulty in stressful compared with nonstressful situations. According to Anderson et al. (2014) solving math problems in normal contexts yields several different functional segments or distinct states that represent different cognitive processes. However, meaningful information pertinent to global neural network stability can also be gleaned from calculating the number of states or state transition frequencies. Under neutral conditions, more states should be indicative of less neural network stability given that HMM defines a unique state when one or more sources (i.e., brain regions) of signal exhibit a unique change in relation to other sources of signal (which, in our case, is power in different brain sources representative of different brain regions). Fewer states should reflect a degree of higher global network stability (or less activity across the brain).

Conversely, past research suggests that completing math tasks in stressful situations engenders a cascade of physiological and performance monitoring processes, including increased recruitment of working memory, emotion regulation, and self-monitoring processes (Schmader et al., 2008). This suggests that multiple networks would be active simultaneously, resulting in decreased brain stability and an increase in brain states, which could tax cognitive resources otherwise needed for optimal performance. To assess the effects of stress on brain network dynamics, it would be critical to examine these processes in individuals who are otherwise comparable on key cognitive dimensions (e.g., do not vary in degree of dispositional attributes such as math anxiety). Stereotype threat,

a situational stressor individuals experience when they fear confirming a negative group stereotype (Schmader, Forbes, Zhang, & Mendes, 2009; Schmader et al., 2008), provides a means of placing one group under stress while holding other aspects of the situation and individuals constant (e.g., solving the same problems).¹ We hypothesized that women under stress (i.e., stereotype threat) should exhibit less global network stability and thus more brain states compared with nonstressed women. We also examined the degree to which state frequency (i.e., global network stability) and state duration degree are key predictors of whether individuals succeed or fail in stressful situations.

METHODS

Participants and Procedures

As part of a larger study (see Forbes & Leitner, 2014), 42 female participants were randomly assigned to either a stereotype threat/diagnostic math test (DMT) condition (21 women) or a control/problem-solving test (PST) condition (21 women). In the DMT condition, participants were told (via a male experimenter) that they would be completing tasks that were diagnostic of their math intelligence. In the PST condition, participants were told (via a female experimenter) that they would be completing tasks that were diagnostic of different types of problem-solving techniques they prefer. To prime the stress response, that is, prime stereotype threat, after the instructions, participants in the DMT condition completed demographic questions and indicated their gender (Steele & Aronson, 1995). Participants in the PST condition completed the same demographic questions except for the question pertaining to gender. These manipulations have been shown to effectively prime stress in numerous studies (e.g., Forbes, Duran, Leitner, & Magerman, 2015). Participants then completed the math task while continuous EEG activity was recorded. All participants were right-handed, native English speakers and had no disabilities that would impair task performance. Two participants were removed from the control group after being identified as outliers on both problem-solving RTs and ventral ACC (vACC) power measures (for the stress manipulation check, see below) via Grubbs test (also known as the extreme studentized deviate method); that is, these individuals had z scores greater than 3.5 SD s below the grand mean ($p < .05$, $M = 9.53$ sec, $SD = 2.12$ sec) of solving time and above the grand mean for vACC power.

Math Task

The math task consisted of more difficult standard multiplication problems (e.g., $19 \times 46 = 874$) and novel, easier problems referred to as sharp problems. Both standard multiplication and sharp problems were inspired by Paynter, Reder, and Kieffaber (2009). Traditional math

problems consisted of standard multiplication problems (e.g., $19 \times 46 = 874$). Sharp problems required a series of steps to solve and always contained the operand “#.” Participants were told how to solve sharp problems and were given several sharp problems to practice beforehand. Specifically, they were told that, when they see the sharp operand (#), they answer the problem by taking the sum of the tens place digits for the two operands, multiplying this number by the sum of the ones place digits, and multiplying this product by 3. For example, $24 \# 16$ would be solved as $(2 + 1) \times (4 + 6) \times 3$, which would be equal to 90. Three answer choices were presented below each problem, and the correct answer was randomly presented in one of the three answer positions. The task consisted of four blocks of 25 problems each for 100 trials/problems (50 hard/multiplication problems and 50 easy/sharp problems). Easy and hard blocks were alternately presented at random. The digits for both easy and hard problems were randomly selected between 10 and 99, and measures were taken to ensure no number overlap between problems. Each trial began with the presentation of a cross hair in the middle of the screen, followed by the problem. Participants were instructed to select their answer on a response box. After each response, participants were presented with feedback on a white background for 2 sec that indicated whether they had answered the problem correct (written in blue) or wrong (written in red). Participants were given 16 sec to solve each problem. Math problem-solving accuracy was calculated by dividing the number of problems solved correctly by the number of problems attempted for hard and easy problems, respectively.

EEG Recording and Preprocessing

Continuous EEG activity was recorded using an ActiveTwo head cap and the ActiveTwo BioSemi system (BioSemi, Amsterdam, The Netherlands). Recordings were collected from 64 Ag–AgCl scalp electrodes. Two electrodes placed below the left eye and to the left of the left eye recorded vertical and horizontal ocular movements, respectively (EOG). During acquisition, the ground electrode was formed by BioSemi’s Common Mode Sense active electrode and the Driven Right Leg passive electrode, as per BioSemi’s proprietary design. EEG activity was digitized with ActiView software (BioSemi) and sampled at 2048 Hz.

In offline analyses, data were referenced by whole-brain average, downsampled postacquisition, and analyzed at 512 Hz. EEG signals were epoched and stimulus locked to 500 msec before participants were presented with a math problem to the time they solved the math problem. Epochs were baseline corrected by subtracting the average value of EEG 100-msec preproblem from the entire epoch. EEG artifacts were removed via FASTER (Fully Automated Statistical Thresholding for EEG artifact Rejection; Nolan, Whelan, & Reilly, 2010), an

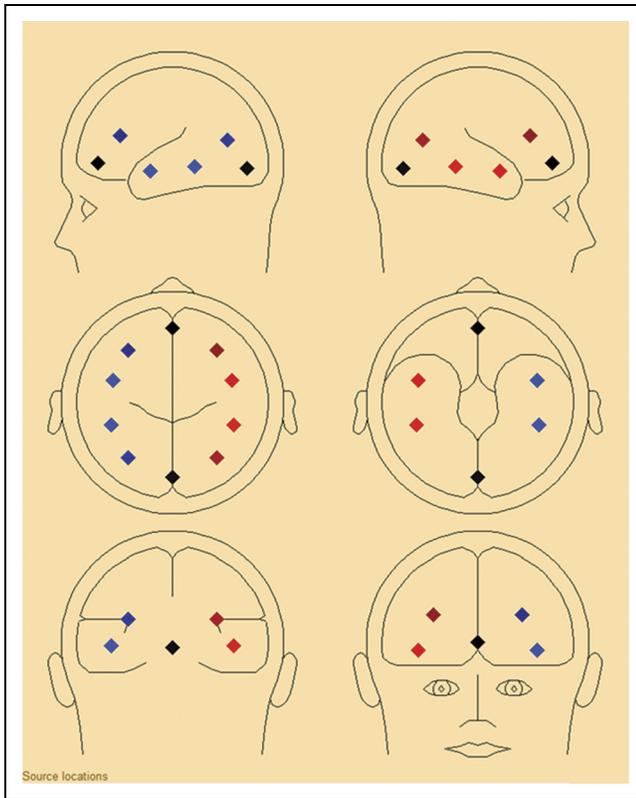


Figure 1. Source model used for HMM analyses. The 10 sources utilized in this study included sources in the frontoparietal network, including bilateral DLPFC (L-RDLPFC), bilateral posterior parietal cortex (L-RIPC), and frontopolar cortex. Sources in occipital cortex, bilateral anterior temporal cortex (L-RATC), and bilateral posterior temporal cortex (L-RPTC) were also included.

automated approach to cleaning EEG data that is based on independent component analysis and multiple statistical steps. Specifically, raw EEG data were initially filtered through a band-pass finite impulse response filter between 0.3 and 55 Hz. First, EEG channels with significant unusual variance (operationalized as activity with an absolute z score larger than 3 SDs from the average), average correlation with all other channels, and Hurst exponent were removed and interpolated from neighboring electrodes using spherical spline interpolation function. Second, EEG signals were epoched and baseline corrected, and epochs with significant unusual amplitude range, variance, and channel deviation were removed. Third, the remaining epochs were transformed through independent component analysis. Independent components with significant unusual correlations with EOG channels, spatial kurtosis, slope in filter band, Hurst exponent, and median gradient were subtracted, and the EEG signal was reconstructed using the remaining independent components. Finally, EEG channels within single epochs that displayed significant unusual variance, median gradient, amplitude range, and channel deviation were removed and interpolated from neighboring electrodes within those same epochs.

Source Montage in Math Solving

FASTER cleaned EEG data were then imported into BESA, transformed into source space, and modeled via 15 equidistant dipole sources across the frontal, central, parietal, and temporal cortices. Sources utilized standard coordinates provided in a standard head model by BESA. Among the 15 dipoles, five of them were excluded because, together, they accounted for less than 5% of the total variance in the total EEG signal. Finally, as gamma waves are believed to play an integral role in information processing and synthesizing information throughout the brain (Buzsáki, 2006), analyses focused on brain activity in the gamma band. Gamma band power (30–55 Hz) over the course of solving a given problem was extracted from each of the 10 sources using short-time Fourier transform (STFT). In addition, the baseline (average value) of each source's gamma activity was subtracted from the entire epoch because we were primarily interested in evoked power activity (Figure 1, Table 1).

HMM

An HMM is a state-space stochastic model. Let s_t and o_t denote assignment of a hidden state and an observation at time t , respectively. Hidden and observable processes over T time length can be denoted by $\{s_1, \dots, s_b, \dots, s_T\}$ and $\{o_1, \dots, o_b, \dots, o_T\}$. For the hidden process, we model it with a first-order Markov chain, which means that the hidden state at time t , that is, s_t , is dependent on the hidden state at time $(t - 1)$, that is, s_{t-1} . The observable variable at time t , that is, o_t , is dependent on the hidden state at the same time, that is, s_t .

Formally, when we consider a number K of hidden states, a hidden process is defined by two probability distributions: state transition probability and initial state

Table 1. Talairach Coordinates of 10 Brain Regions Utilized in This Study

	x	y	z
L-anterior temporal	-44.7	5.4	5.3
L-posterior temporal	-44.6	-37.3	-4.0
L-DLPFC	33.8	32.7	25.8
L-posterior parietal	-33.8	-69.4	15.9
Frontopolar	0	55.1	4.7
Occipital	0	-86.6	-9.0
R-DLPFC	33.8	32.7	25.8
R-posterior parietal	33.8	-69.4	15.9
R-anterior temporal	44.7	5.4	5.3
R-posterior temporal	44.6	-37.3	-4.0

L = left; R = right.

probability. A state transition probability $A = [a_{ij}]_{i,j \in \{1, \dots, K\}}$, where $a_{ij} = P(s_t = j | s_{t-1} = i)$, denotes the probability of changing from one hidden state ($s_{t-1} = i$) to another hidden state ($s_t = j$). When $i = j$, suggesting the state does not transition, the state transition probability is replaced by a duration probability $D = [P_i(t - \tau)]$ for a specific hidden state, where τ is the initial time for the current specific hidden state and $t - \tau$ represents the duration of the current state. In this work, we use a Gaussian distribution for duration probability. In addition, an initial state probability $\Pi = [\pi_i]_{i \in \{1, \dots, K\}}$, where $\pi_i = P(s_1 = i)$ models the probability of starting from a specific hidden state at time $t = 1$, denotes where the state transition begins. The observable process is depicted by an emission probability density function $B = [b_i]_{i \in \{1, \dots, K\}}$, where $b_i = P(o_t = x_t | s_t = i)$ denotes the likelihood of observing the specific observation x_t when residing at the hidden state of i . In this work, we use a single Gaussian distribution for emission probability. Thus, an HMM is completely defined by the parameter set of $\lambda = (\Pi, A, B, D)$ (Suk, Wee, Lee, & Shen, 2016; Rabiner, 1989).

HMM Analyses

To construct HMM models, first, a learning step was applied where an individual HMM model was optimized for best fitting all the data in each condition. This model was trained via an efficient method proposed by Yu and Kobayashi (2003, 2006). HMM analyses were conducted on periods during which individuals solved math problems (only when the problems were present)

using the evoked EEG gamma power curve elicited from the 10 cortical regions outlined in our source model. Once the HMM models were set up, the second step of decoding was applied to the gamma curves stemming from each source to identify what is the most likely state sequence in the model that produced the observations. We obtained a series of states generated from the gamma curve in every participant and utilized state count to index brain network stability. Figure 2 illustrates the flowchart of HMM analysis in our study.

Condition Comparison

All trials were partitioned into a 2 (condition: stress vs. control) \times 2 (problem difficulty: easy vs. hard) \times 2 (accuracy: correct vs. wrong) matrix, and eight HMM models were established to model the problem-solving process.

Trials vs. Average

We incorporated an average scheme for HMM analyses, which utilized participant-by-participant evaluation as opposed to trial by trial to guarantee reliably differential patterns. A practical problem facing the average approach is that the time intervals of all trials are different. Nevertheless, we reasoned that every participant engages in similar cognitive mechanisms or stages during the problem-solving process. As such, we applied a time standardization approach. We interpolated the time series of each trial using the cubic spline method to make all trials of the same duration. Given that participants were provided 16 sec

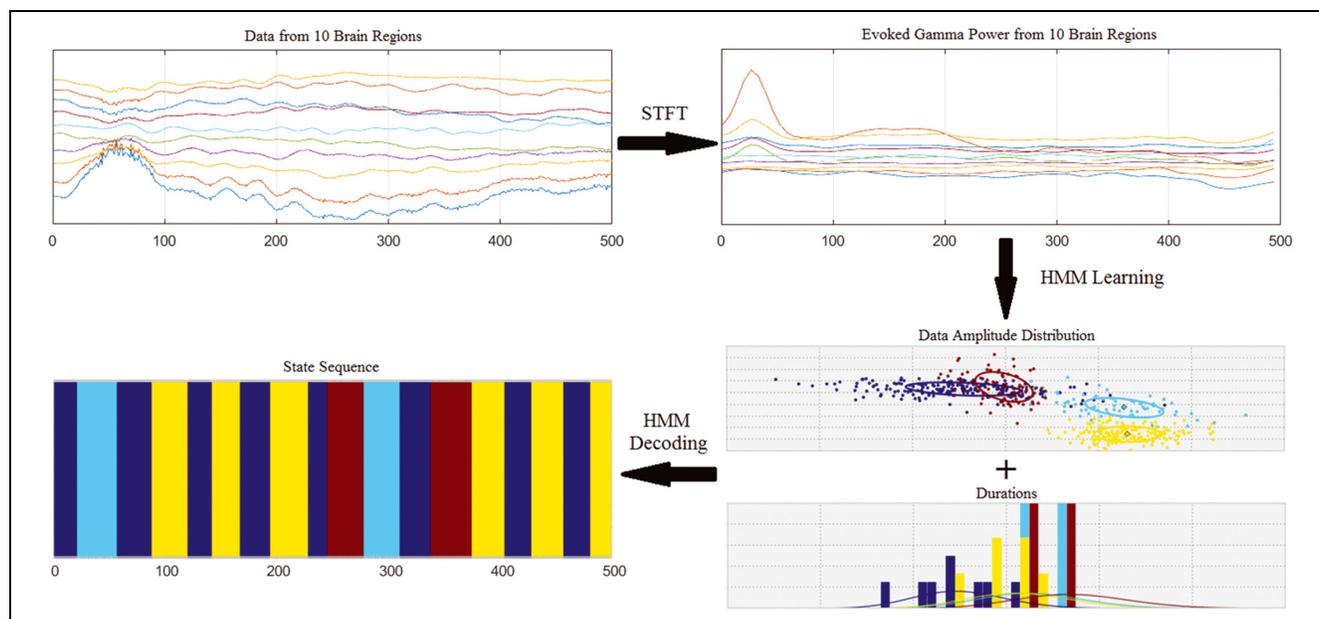


Figure 2. Flowchart for HMM analyses. Time series were extracted from 10 brain regions outlined in the source model. STFT was then applied to obtain the gamma power curve. HMM models were trained using all trials of data in each condition. Gamma curves from the 10 sources were embedded into a 10-dimensional data space (to better illustrate, we present only two dimensions here) and were considered to conform to a specific 10-dimensional Gaussian distribution in each state. The duration probability of each hidden state was assumed to satisfy a Gaussian distribution. The trained HMM model was then utilized to decode each trial to identify a specific sequence of states.

to solve each math problem, we interpolated all trials to 16 sec in length. Because gamma power after the STFT does not tend to vary rapidly, we downsampled gamma power by averaging 100 temporal points into one time point. This allowed us to circumvent computational burdens associated with the complex HMM analyses employed in this study while still preserving the original features of the data.

Determination of Hidden State Number

Given that the likelihood of fit always increases in accordance with the number of states added to the model in the training set, theoretically, one could approach 100% model fit with a large number of states. However, an excessive number of states could yield a model that overfits the data. The number of states in this study was determined via a threshold for marginal increase of recognition rate for the whole data set (Celeux, 2007). That is, the number of state n was selected when the recognition rate increased less than 1% with respect to the number of state $n + 1$. On the basis of these criteria, we selected 6 as the optimal number of states to construct the model.

Stress Manipulation Check in Neural Data

Given the neural-specific nature of our hypotheses and the disconnect between self-report and biological reactions often evident in stereotype threatening contexts (Schmader et al., 2008), where threatened individuals may exhibit biological patterns consistent with anxiety and stress responses but fail to report explicit perceptions consistent with stress, we examined specific neural measures to provide support for whether the stereotype threat manipulation successfully engendered neural responses normally evident in stressful situations in the stereotype threat condition. To do this, we first conducted a meta-analysis using the Neurosynth software (neurosynth.org; Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). Neurosynth collects thousands of published articles reporting the results of fMRI studies and conducts forward and reverse inference meta-analyses. Reverse inference provides a stronger estimate of the relative specificity with which a particular pattern of brain activity is associated with a specific cognitive process; thus, the results from the reverse inference analysis were applied in this study. Although the focus of this study was on “stress,” stereotype threat has been associated with similar emotional constructs including anxiety, fear, and threat. Thus, we included the terms “stress,” “anxiety,” “fear,” “threat,” “pain,” and “negative emotion” in the meta-analysis. These analyses yielded activation clusters across all searched terms specifically in vACC. Thus, we selected vACC as the brain region most likely to provide a reliable estimate of our stress manipulation. To further verify that vACC activation is more strongly associated with

stress as opposed to other cognitive processes that may be activated in our study, such as problem-solving processes or working memory, we performed effect size-wise meta-analytic contrasts interactively between emotional terms (stress, anxiety, and fear) and cognitive terms (executive, working memory, and attention) using Neurosynth (Lieberman, Burns, Torre, & Eisenberger, 2016). These analyses repeatedly yielded significant contrasts in vACC (Figure 3), meaning we identified voxels in vACC in which the average likelihood of activation reported was larger for studies containing the emotion-based keywords compared with studies containing the cognitive-based keywords.

To measure vACC power, we used a distributed source localization system. All a priori sources used were identified and calculated via forward and inverse models utilized by MNE-python (Gramfort et al., 2013, 2014). Cortical surfaces extracted with FreeSurfer were subsampled to approximately 10,240 equally spaced vertices on each hemisphere, which is constrained by the default average template of anatomical Montreal Neurological Institute MRI. The cortical surface was divided into 68 anatomical ROIs (34 in each hemisphere) based on the Desikan–Killiany atlas (Desikan et al., 2006), and the power from seed voxels from each area was collapsed to calculate the total power in that area. vACC power was measured using power within specific ROIs where our target coordinates found from Neurosynth were included. Given that the minimum problem-solving time for a single trial found in this study was 1.39 sec, power analyses focused on epochs containing the first 1000 msec of problem solving for each problem type; this ensured that all trials for each participant would be represented and included in analyses.

We also extracted power from a region that appeared to be significantly less involved in stress-oriented processing. Analyses in Figure 3 repeatedly yielded significant contrasts in SMA, meaning voxels in SMA were identified in which the average likelihood of activation reported was larger for studies containing the cognitive-based keywords compared with studies containing the emotion-based keywords. Thus, SMA was not expected to vary in power as a function of increased stress during problem solving and thus condition.

RESULTS

Stress Manipulation Check

A 2 (Condition: stress or control) \times 2 (Problem difficulty: easier or harder) \times 2 (Accuracy: incorrect or correct) mixed factors ANOVA with repeated measures on the latter two variables was conducted on vACC power within gamma frequency bands in the first 1000 msec of problem solving. Results revealed a main effect for condition, $F(1, 39) = 4.82, p = .034, \eta^2 = 0.11$, indicating that women in the stress condition ($M = 0.859, SD = 0.105$)

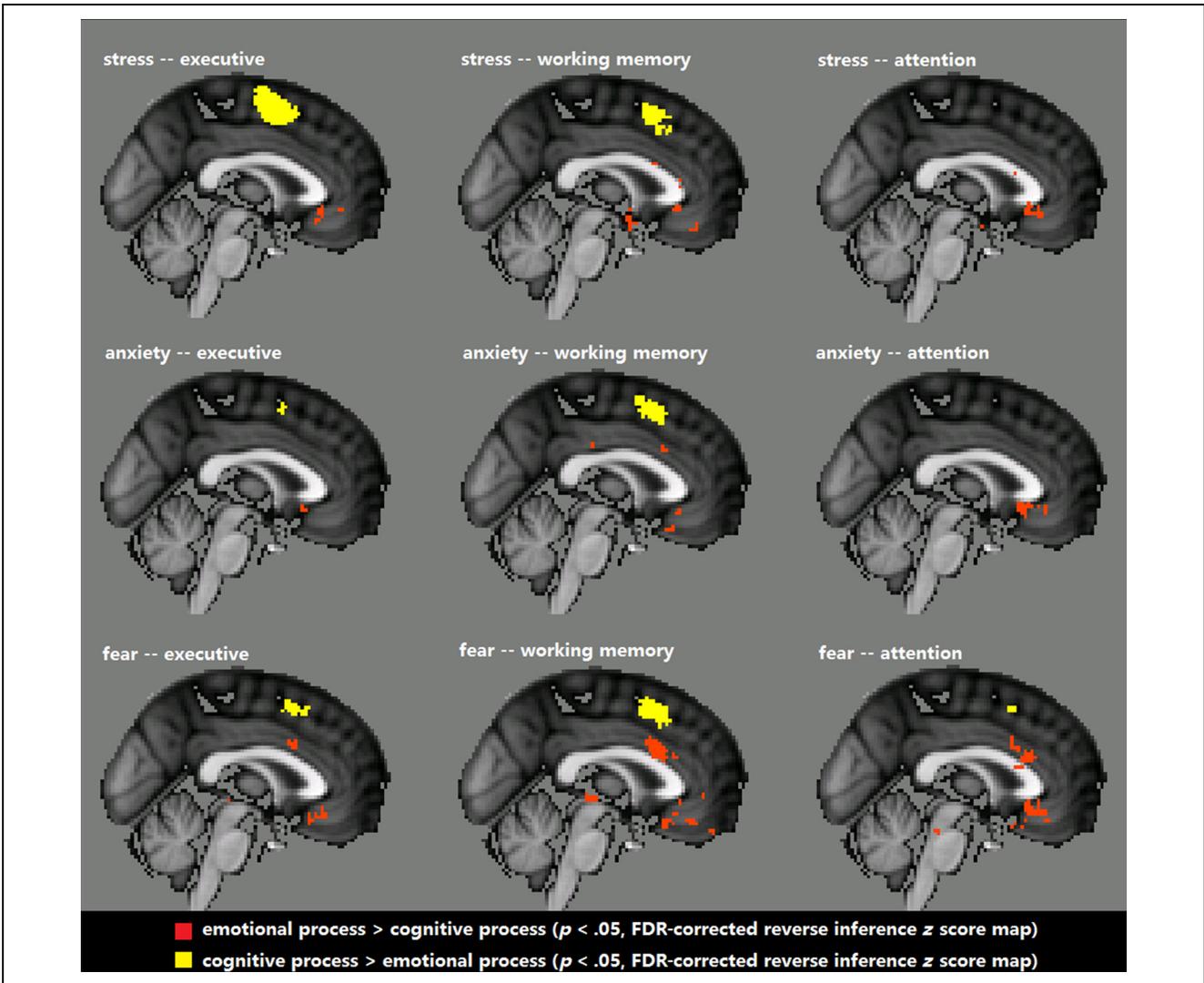


Figure 3. Selection of neural regions for stress manipulation check. Neural region selected for stereotype threat manipulation check was determined via the software Neurosynth (neurosynth.org; Yarkoni et al., 2011). Effect size-wise meta-analytic contrasts were performed interactively between emotional terms (stress, anxiety, and fear) and cognitive terms (executive, working memory, and attention) using Neurosynth-python. These analyses repeatedly yielded significant contrasts in vACC, meaning we identified voxels in vACC in which the average likelihood of activation reported was larger for studies containing the emotion-based keywords compared with studies containing the cognitive-based keywords. FDR = false discovery rate.

exhibited higher vACC gamma power compared with women in the control condition ($M = 0.521$, $SD = 0.113$). No other main or interaction effects were significant ($ps > .12$), suggesting that the stress manipulation elicited neural activity consistent with a stress response

among women in the stress condition. Results of additional ANOVAs conducted on the three other frequency bands are presented in Table 2.²

In an effort to document discriminant validity for vACC power, it is important to demonstrate that activation in

Table 2. Statistical Results for Stress Manipulation Check

Frequency	<i>df</i> ₁	<i>df</i> ₂	Mean Power (Control)	Mean Power (Stress)	<i>F</i>	<i>p</i>
Theta	1	39	1.87	2.75	6.859	.012
Alpha	1	39	1.21	1.60	1.907	.175
Beta	1	39	0.78	1.23	4.788	.035
Gamma	1	39	0.52	0.86	4.000	.034

regions not expected to be affected by stress does not vary as a function of condition. Analyses reported in Figure 3 repeatedly yielded significant contrasts in SMA, suggesting that the average likelihood of activation reported in SMA voxels was larger for studies containing the cognitive-based keywords compared with studies containing the emotion-based keywords. Thus, we hypothesized that power in SMA would not increase as a function of increased stress during the problem solving process. To test this, we conducted a 2 (condition: stress or control) \times 2 (problem difficulty: easier or harder) \times 2 (accuracy: incorrect or correct) mixed factors ANOVA with repeated measures on the latter two variables on SMA power within the gamma frequency band on the first 1000 msec of problem solving. Results indicated that there was no main effect for condition ($p = .69$) and there were no any interaction effects ($ps > .25$), suggesting that cognitive processes ascribed to SMA activation (e.g., planned motor movements) do not vary among women in either condition.

Solving Time and Accuracy on Math Task

A 2 (Condition: stress or control) \times 2 (Problem difficulty: easier or harder) \times 2 (Accuracy: incorrect or correct) mixed factors ANOVA with repeated measures on the latter two variables was conducted on average problem-solving time. Results revealed a main effect for Condition, $F(1, 39) = 12.15, p = .006, \eta^2 = 0.23$, indicating that stressed women ($M = 8.93, SD = 0.33$) were faster at solving all problem types compared with women in the control condition ($M = 10.32, SD = 0.35$). There was also a main effect for Accuracy indicating that women were faster on problems solved correctly ($M = 8.84, SD = 0.22$) compared with problems solved incorrectly ($M = 10.41, SD = 0.29$), $F(1, 39) = 23.35, p < .001, \eta^2 = 0.41$. There was also a two-way interaction between Accuracy and Problem difficulty, $F(1, 39) = 20.13, p < .001, \eta^2 = 0.34$. Simple effect analyses indicated that easier problems solved correctly ($M = 8.47, SD = 0.26$) were solved faster than harder problems solved correctly ($M = 9.20, SD = 0.29$), $F(1, 39) = 5.08, p = .03, \eta^2 = 0.12$. Simple effects were not found in incorrectly solved problems. No other main or interaction effects were significant ($ps > .21$).

Another 2 (Condition: stress or control) \times 2 (Problem difficulty: easier or harder) mixed factors ANOVA with repeated measures on the latter variable was conducted on average problem-solving accuracy. These results indicated that easy ($M = 0.78, SD = 0.20$) math problems were solved more accurately compared with hard ($M = 0.48, SD = 0.20$) math problems, $F(1, 39) = 197.40, p < .001, \eta^2 = 0.84$. No other effects were significant, $ps > .77$.

Together, findings indicate that stressed women were faster at solving all problem types compared with women in the control condition, which is consistent with past research suggesting that stress prompts avoidance behaviors that motivates individuals to solve problems

as quickly as possible in an attempt to conclude the stress-inducing task (Ashcraft, 2002). However, this avoidance is usually accompanied by a speed-accuracy trade-off (Ashcraft, 2002), which was not found in this study. Actually, many studies (e.g., Forbes et al., 2015; Spencer, Steele, & Quinn, 1999) tend to show an overall main effect of reduced performance for participants under stress. It is also possible, however, that individual variation in this study was more robust given the length of the math task, which was much longer than typical stereotype threat studies.

To account for this, we analyzed performance data using a more sensitive analytic approach that more accurately accounted for individual variability across time. Performance data were analyzed using generalized estimating equations (GEEs; Ballinger, 2004; Liang & Zeger, 1986), specifying a binary logit model given that the outcome variable is binary (incorrect = 0, correct = 1) and an unstructured working correlation matrix (Fitzmaurice, Laird, & Rotnitzky, 1993). Mimicking the ANOVAs, we initially included the main effects of Condition and Problem difficulty in addition to the Condition \times Problem difficulty interaction as predictors of performance. These analyses only yielded a main effect for Problem difficulty, $B = -3.186$ ($SE = 1.0270$), Wald $\chi^2(1) = 9.627, p < .01$, indicating that, for every 1-unit change in problem difficulty (from easy, coded 0, to hard, coded 1), the log odds of getting the question correct decreased by 3.186 units (all other $ps > .26$). However, given the conflicting patterns found across our data specific to problem difficulty (e.g., there was no main effect for Problem type in vACC, RT, or some state analyses reported below), the nature of our harder problems (traditional multiplication problems), and the well-documented finding that stereotype threat performance effects are specific to very challenging problems that often take the form of more complex, working-memory-intensive word problems typical to those found on the math portion of the SAT and GRE (and indeed actually facilitate performance on easier problems; Ben-Zeev, Fein, & Inzlicht, 2005; Spencer et al., 1999), we conducted an additional GEE analysis that collapsed across problem difficulty and included only the main effect of Condition as a predictor. This analysis revealed a main effect for Condition on performance, $B = 0.18$ ($SE = 0.09$), Wald $\chi^2(1) = 4.10, p < .05$, indicating that, for every 1-unit change in condition (from stress, coded 0, to control, coded 1), the log odds of getting the question correct increased by 0.182 unit. That is, women in the control condition were more likely to get the question correct than women in the stress condition, providing evidence for a typical stereotype-threat-based stress effect on performance.

It should also be noted that Forbes and Leitner (2014) found that, compared with participants in the control condition, participants in the stress condition (the same participants used in this study) underperformed on difficult math problems to the extent that they were more

identified with math, which is a typical stereotype threat effect found in the literature (Spencer et al., 1999). Coupled with the vACC and GEE findings, these results indicate that the stress manipulation was mostly successful.

State Count Analysis

HMM can be linear or fully connected. Whereas fully connected refers to the process where every state of the model can be reached from every other state of the model, linear connected means a particular state can only be reached from another specific state of the model. A fully connected HMM model creates the possibility that not all states appear in one single EEG trial; however, one state can appear more than one time in one single EEG trial. Conversely, a linear connected HMM is constrained by the number of states identified by the model definition on every single EEG trial (which in our model is six). In our study, we utilized the fully connected model because, psychologically, it is reasonable that the brain functionally switches its state back and forth during a particular cognitive process. In addition, it theoretically provides a more veridical index of global neural network stability. In this case, we can have an accurate state count for any given trial, which then allows the state count to represent global network stability (where higher state counts would equate to less global network stability).

To determine whether the number of states played a role in successful problem solving as a function of stress, a 2 (Condition) \times 2 (Problem difficulty) \times 2 (Accuracy) mixed factors ANOVA with repeated measures on the latter two variables was conducted on state count. Results yielded a main effect for Condition where women under stress ($M = 7.09$, $SD = 0.35$) exhibited more states than women not under stress ($M = 5.87$, $SD = 0.37$), $F(1, 39) = 5.74$, $p = .02$, $\eta^2 = 0.13$, suggesting that stressed individuals exhibited more global neural network instability compared with nonstressed individuals (Figure 4C). In addition, a marginal effect for Problem difficulty was found indicating that hard problems elicited more states ($M = 6.78$, $SD = 0.33$) than easy problems ($M = 6.18$, $SD = 0.27$), $F(1, 39) = 3.83$, $p = .057$, $\eta^2 = 0.09$. In addition, there was a two-way interaction between Accuracy and Condition, $F(1, 39) = 7.72$, $p = .008$, $\eta^2 = 0.17$. Simple effect analyses indicated that the control group exhibited more states ($M = 6.40$, $SD = 0.47$) when they correctly solved problems compared with when they incorrectly solved problems ($M = 5.34$, $SD = 0.44$), $F = 4.37$, $p < .05$, $\eta^2 = 0.10$. Conversely, stressed women tended to exhibit fewer states ($M = 6.64$, $SD = 0.43$) when they correctly solved problems compared with when they incorrectly solved problems ($M = 7.55$, $SD = 0.41$), $F = 3.77$, $p = .06$, $\eta^2 = 0.08$. Furthermore, whereas stressed women exhibited comparable state counts compared with women in the control condition when problems were solved correctly, on incorrect problems, stressed

women exhibited far more state counts compared with women in the control condition ($F = 13.27$, $p < .001$, $\eta^2 = 0.25$; Figure 4C). No other main or interaction effects were significant ($ps > .77$).

To explore whether state count had any relationship with math performance, a multiple regression analysis was conducted on problem-solving accuracy and state count as a function of the factors that were found to interact in our state count ANOVA analyses. The following factors were included in the multiple regression analysis: Condition, State count for wrong problems, State count for correct problems, and their interaction terms (Condition \times Correct, Condition \times Wrong, Correct \times Wrong, and Condition \times Correct \times Wrong). Results indicated that none of the factors yielded a significant effect, $ps > .23$.

Although these initial analyses indicated that there were no linear relationships among the variables, it is also possible that nonlinear relationships existed between state count and math performance. Given consistencies among women in both conditions between state count and problems solved correctly, one possibility is that an optimal state count exists for successful problem solving. That is, there is an optimal level of global network stability in relation to performance that exists regardless of whether an individual is stressed, not unlike Yerkes–Dodson’s law of performance that posits a quadratic relationship between arousal and performance, where too little or too much arousal can compromise performance on a given task. To test this idea, a second-degree polynomial regression analysis was conducted on average state count. Given that, in our initial ANOVA, we only found a two-way interaction between Condition and Accuracy but not Problem difficulty (a separate ANOVA conducted on easy and hard yielded no significant interactions, $ps > .12$), we regressed math accuracy (collapsing across easy and hard problem accuracy) on to average state count (collapsing across the easy–correct [EC], easy–wrong [EW], hard–correct [HC], and hard–wrong [HW] conditions). Results yielded a significant quadratic effect, $F(2, 37) = 4.36$, $p = .02$, $\eta^2 = 0.19$ (Figure 4B),³ suggesting that there was an optimal level of states associated with math accuracy;⁴ too much or too little global network stability was associated with suboptimal performance. Quadratic regression analyses conducted on math accuracy and state count for easy and hard problems separately also noted a trend for a quadratic effect in both problem difficulties; however, these effects were not significant ($ps < .15$).

U-shape Relationship Test

Given recommendations by Simonsohn (2016), it would also be important to test whether the inverse-U-shape pattern found in our data observes significant monotonic increases and decreases at the two ends of the curve. To do this, a two-line test was conducted on the secondary

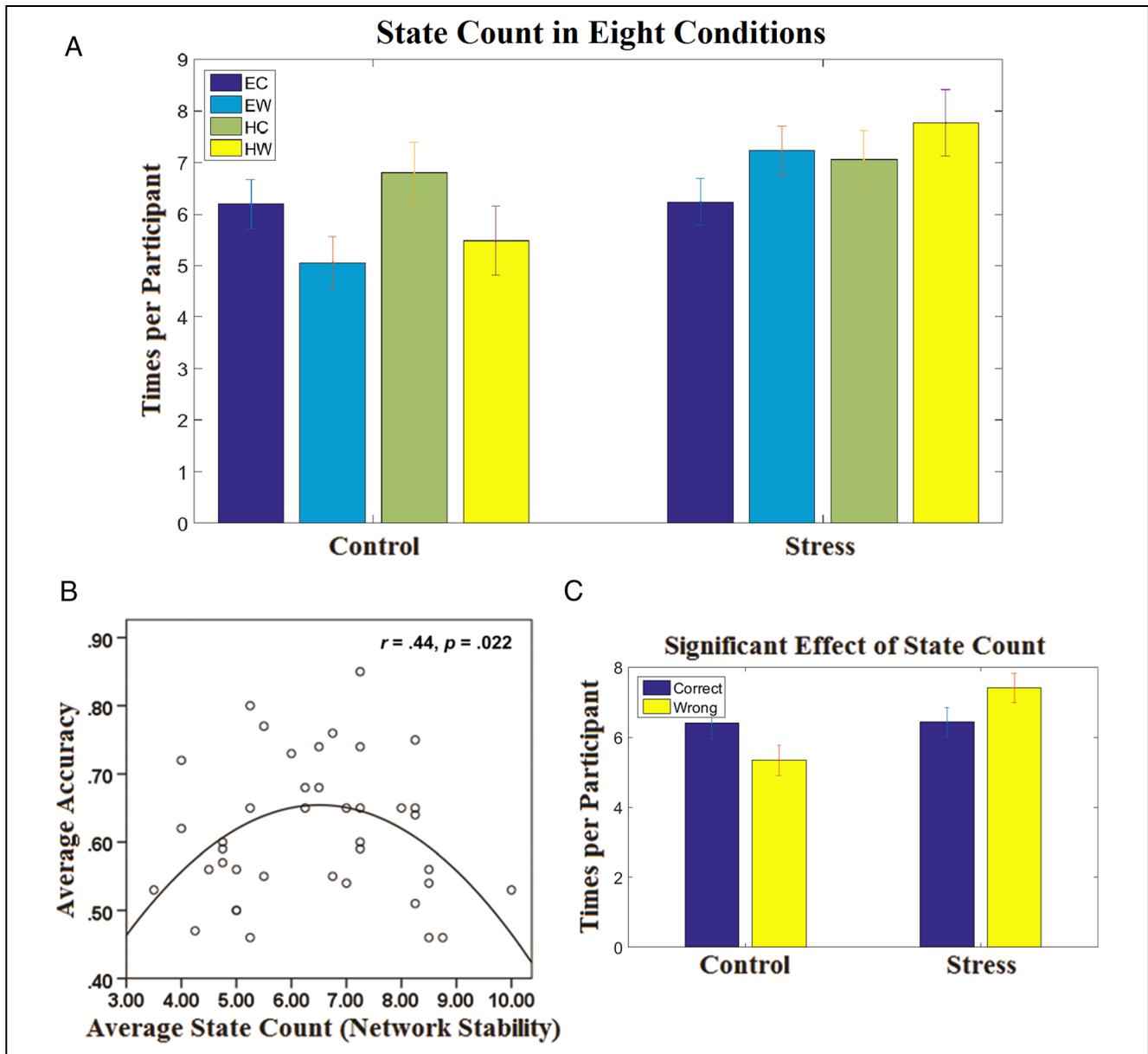


Figure 4. Global neural network stability (brain state count) as a function of condition and performance. (A) Brain state count as a function of the eight conditions: It is clear that women under stress exhibited more states than women not under stress overall. (B) Math solving accuracy increases with global neural network stability, that is, state count, but only up to a point. When stability is either too low or high, performance decreases. (C) Significant effect of state count: In women in control, they have more states when they correctly solve the math problems, whereas in women under stress, they have fewer states when they correctly solve the math problems ($p = .008$). However, when women correctly solved the math problems, there appeared little differences in state count.

polynomial curve found between number of states and solving accuracy. The adjusted breakpoint was found at a state number of 6.87. Two separate linear regression analyses were then conducted on two data sets separated by the adjusted breakpoints. Results indicated that both the monotonic increase curve, $t(22) = 2.138$, $p = .044$, $\beta = 0.423$, $r = .423$, and monotonic decrease curve, $t(16) = -2.097$, $p = .046$, $\beta = -0.471$, $r = .471$, were significant, suggesting that the U-shape relationship found in our analyses was valid.

State Count and Solving Time

The solving time in our study has been standardized; nevertheless, one might argue that greater state counts stem from problems that take longer to solve (which would suggest that harder problems would have more state counts, as would women in the control condition who took longer to solve all problems on average). To alleviate this concern, within-cell correlations were conducted on solving time and state count in both the

control and stressed conditions. Results indicated that neither correlation was significant ($r_{\text{control}} = -.008, p = .72; r_{\text{stress}} = .197, p = .28$), indicating that state count was not associated with actual solving time.

State Identification

States were defined with HMM via gamma power extracted from each of the 10 sources in this study. Consistent with past research (Anderson et al., 2014), we operationalized several reliable patterns of neural activity present in the HMM models. One might assume that, with eight HMM models, there should be different state definitions unique to each model; that is, with eight models and six states per model, there could be 48 distinct states identified. However, some of the states across conditions had identical power distributions in 10 brain regions, suggesting that they could represent a unique, meaningful functional state.

To avoid ambiguity that may arise from confounding similar states, as well as to simplify our analyses, we categorized the 48 states into several similar upper-level states. To do this, a k -means cluster approach, with correlation distance function, was conducted on the power distribution from 10 brain regions to group the 48 states. The k -mean solution was evaluated across values of k ranging from 6 to 20 using the silhouette metric. The k value that returned the highest silhouette score was 9. Of the nine states, one was found to be unique to a specific condition and was thus merged with another state with the closest distance. Thus, our final analyses focused on eight states. Of the eight states, four states appeared in all conditions, two states were only present in the control condition, and two states only appeared in the stress condition. Below is a brief description of the identified states with respect to the gamma power of each neural source across the brain.

The four states below were found across all conditions:

State 0 = extremely low global gamma power: This state is characterized by extremely low power across all sources. This state, which appears at the beginning and end of the math solving process, is a product of signal processing procedures. Thus, it was removed from data representation in all the figures below as well as the state count analysis previously.

State 1 = temporal cortical dominant, moderate frontal, and occipital gamma power: This state is characterized by one or more of the temporal cortical sources exhibiting extremely high power compared with other sources in the same state and temporal power in other states. Gamma power in prefrontal regions is also fairly high but less so compared with the temporal cortex.

State 2 = moderate gamma power: No brain region exhibited dominant activity in gamma power, and the sources exhibited moderate gamma power.

State 3 = occipital dominant, weak frontopolar gamma power: This state is characterized by dominant power in occipital cortex, accompanied by extremely low activity in the frontopolar source.

The two states below were found in the control group only:

State 4 = high global gamma power, frontopolar dominant: This state is characterized by high gamma power globally, especially in the frontopolar source.

State 5 = low global gamma power: This state is characterized by low gamma power globally. All brain sources exhibited low levels of gamma power, and no region exhibited higher levels of gamma power compared with others.

The two states below were found in the stress group only:

State 6 = average global gamma power across sources: This state is characterized by average power globally (0 for only evoked component left) across all sources. None of the brain regions exhibited stronger power or fluctuations.

State 7 = moderate global gamma power: This state is characterized by slightly higher power than State 6 in all brain regions.

Four typical examples illustrating state distributions and duration in relation to performance were selected from participants in the stress and control conditions are provided in Figure 5.

Anderson et al. (2014) speculated about state functions by viewing activity in specific brain regions and states and inferring what cognitive processes might have been occurring in relation to known functions of a given region. Instead of doing that, we constrained our state descriptions to broader representations of the network, because it is difficult to infer what exactly these states mean without engaging in reverse inference. This is particularly true given that EEG has lower spatial resolution and gamma power could reflect activity stemming from subregions within a broader anatomical region (Table 3).

Multiple Regression Analyses

Although it is difficult to infer what exactly each state's function is based on activity in a given brain region, we can still identify key states with respect to performance by examining how the duration of each state contributes to problem-solving accuracy. Separate multiple regression analyses were conducted on problem-solving accuracy (collapsing across EC, EW, HC, and HW groups) and duration of each state in stressful and nonstressful situations. A stepwise approach was utilized to identify the states that were the strongest predictors of problem-solving accuracy while still accounting for all states. Results indicated that solving accuracy in the control condition was

Figure 5. Prototypical example products of HMM analyses conducted on problem-solving intervals. All examples are with respect to solving hard math problems, and the start/end state is removed. (A) A problem solved correctly by a woman in the control condition. (B) A problem solved incorrectly by a woman in the control condition. (C) A problem solved correctly by a woman under stress. (D) A problem solved incorrectly by a woman under stress. Note the dominance of temporal regions in State 1 and occipital cortex in State 3 and that more states were evident when women were under stress compared with when women were in the control condition, particularly when problems were solved incorrectly.

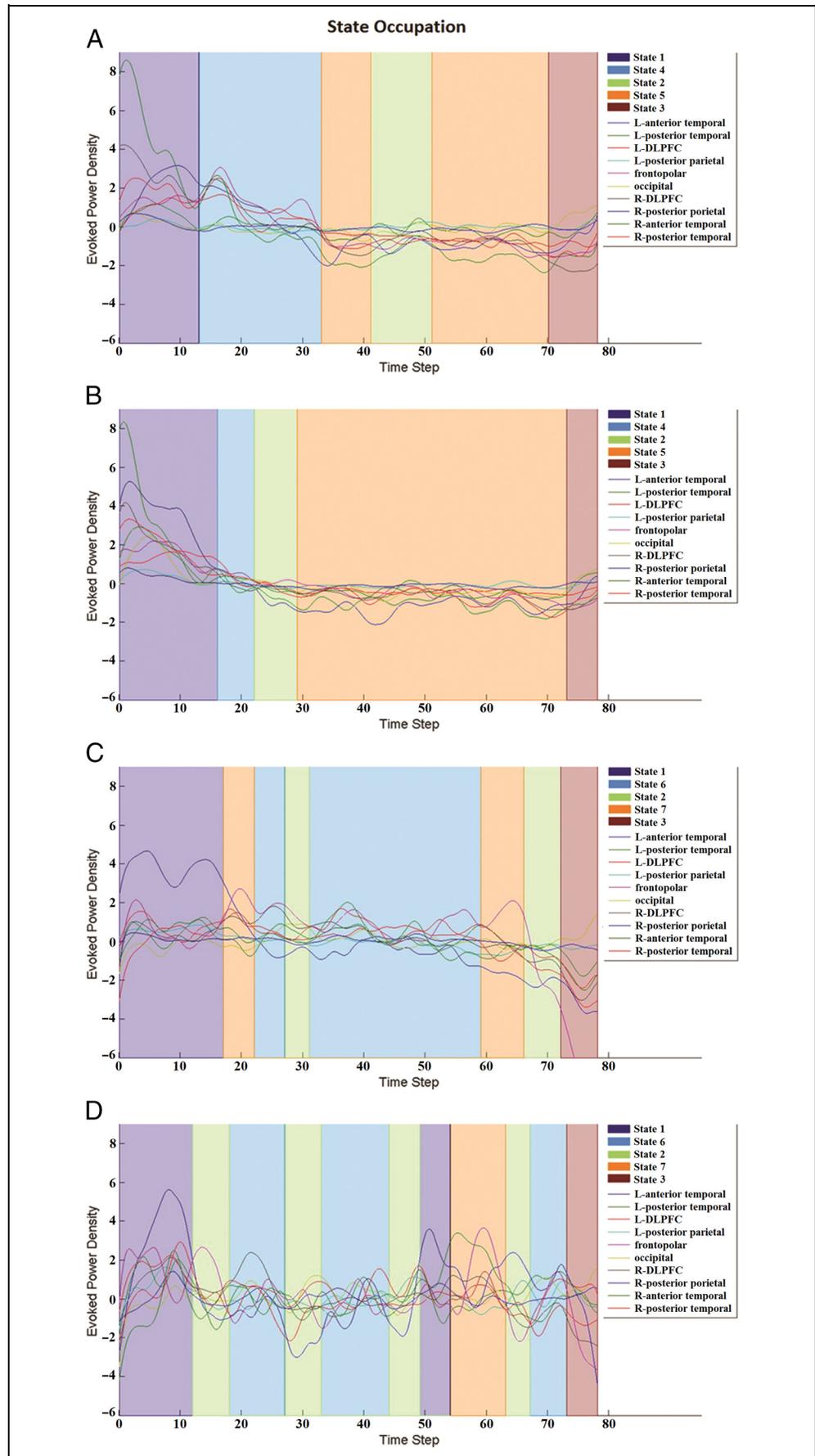


Table 3. Average Power of Brain Regions with Respect to All Seven Isolated States

	<i>State 1</i>	<i>State 2</i>	<i>State 3</i>	<i>State 4</i>	<i>State 5</i>	<i>State 6</i>	<i>State 7</i>
L-anterior temporal	1.62	-0.27	-0.19	0.09	-0.60	-0.05	0.25
L-posterior temporal	0.91	-0.12	0.13	0.01	-0.34	0.03	0.11
L-DLPFC	0.80	-0.07	-0.38	0.29	-0.26	0.03	0.27
L-posterior parietal	0.25	-0.03	0.28	-0.04	-0.17	-0.01	-0.01
Frontopolar	0.93	0.10	-0.54	0.63	-0.27	0.02	0.34
Occipital	0.30	-0.05	0.56	-0.01	-0.25	-0.06	0.05
R-DLPFC	0.81	-0.11	-0.26	0.28	-0.33	0.09	0.17
R-posterior parietal	0.16	-0.06	0.30	-0.10	-0.18	-0.02	0.01
R-anterior temporal	1.76	-0.33	-0.20	0.21	-0.97	-0.28	0.25
R-posterior temporal	0.73	-0.14	0.29	0.03	-0.47	-0.13	0.13

higher to the extent that participants exhibited longer State 4 durations, $t(20) = 2.117, p = .049, \beta = 0.457, r = .434$. Conversely, solving accuracy in the stress condition was more dependent on the duration of State 6, $t(18) = 2.103, p = .049, \beta = 0.434, r = .457$. Thus, better solving accuracy in the control condition was associated with longer durations in a state that consisted of higher gamma power in frontopolar cortex, a brain region integral for higher-level cognitive processes and problem solving (Anderson, Betts, Ferris, & Fincham, 2011; Braver & Bongiolatti, 2002), and higher power in bilateral DLPFC, a brain region integral for executive function and working memory processes (Rottschy et al., 2012). Conversely, better solving accuracy in stressful situations was associated with longer durations in a state that consisted of reduced global gamma power.

States 4 and 6 Represent Global Synchrony for Control and Stress Conditions, Respectively

To gain further insight into the role of global synchrony in the control and stress conditions, we measured whether

oscillatory activity in each source varied in conjunction with one another or oscillated more randomly. We indexed global synchrony by averaging the Pearson's correlation coefficient between all possible pairwise sources in a specific state; for example, we averaged the correlations between the time-variant power curve in gamma band that exhibited a specific state in Sources 1 and 2, 1 and 3, 1 and 4, and so forth within States 1–8 in the eight conditions (control vs. stress, easy/hard, wrong/correct). Findings indicated that, interestingly, States 4 and 6 were found to interact globally (i.e., sources were more correlated with each other) better than any other state in the control and stress groups, respectively (see Table 4). That is, within the eight conditions, the 10 sources were more strongly correlated with each other in those two states specifically. In conjunction with the regression analyses on state duration, our findings suggest that better solving accuracy was associated with longer durations in states that consisted of sources that better interacted with one another globally. Particularly, this is consistent in both control (State 4) and stressful (State 6) situations. This is consistent with

Table 4. Average Values of Correlation Coefficient among All Pairs of Brain Regions with Respect to Each State in All Eight Conditions

	<i>Control</i>				<i>Stress</i>			
	<i>EC</i>	<i>EW</i>	<i>HC</i>	<i>HW</i>	<i>EC</i>	<i>EW</i>	<i>HC</i>	<i>HW</i>
State 1: temporal dominant	.43	.25	.30	.42	.21	.42	.31	.30
State 2: partially moderate	.39	.33	.31	.39	.22	.21	.34	.23
State 3: occipital dominant	.41	.37	.43	.38	.42	.44	.42	.32
State 4: frontopolar dominant	.50	.47	.67	.57				
State 5: global low	.34	.36	.43	.42				
State 6: global average					.58	.53	.49	.54
State 7: global moderate					.32	.15	.36	.31

past research that indicates that global connectivity, especially connectivity between frontoparietal regions, increases during various cognitive tasks (Vakhtin et al., 2014; Rosenberg-Lee et al., 2011; Sauseng et al., 2005).

DISCUSSION

In this study, stress-based differences in performance were found with respect to a novel means for measuring neural network stability during the problem-solving process: HMM. When individuals solved problems in a stressful situation, they exhibited more brain states overall, suggesting that global neural networks were less stable (i.e., oscillations in power in sources across the brain with respect to another). This was particularly evident in brain regions known to play an integral role in the problem-solving process (frontoparietal neural network regions) and when stressed women incorrectly solved math problems. Conversely, nonstressed women exhibited fewer states when they incorrectly solved problems. Additional post hoc analyses indicated that a quadratic relationship existed among all participants with respect to state count and performance, not unlike the Yerkes–Dodson law that describes the quadratic relationship between arousal on performance. Specifically, all women appeared to perform better on the math task (both easier and harder problems) when they exhibited state counts that were neither too few nor too many. This provides evidence for an optimal level of global network stability that is associated with optimal performance, regardless of the context.

When examining the nature of the states themselves, results provide evidence that the neural mechanisms utilized to successfully solve problems are different among stressed compared with nonstressed individuals. Regression analyses conducted on performance accuracy in relation to the amount of time individuals spent in the six states indicated that nonstressed individuals performed better on all problems to the extent that their brains spent more time in a state defined by increased power in frontopolar cortex and bilateral DLPFC, regions that are integral for executive function, working memory, and problem solving in general (Rottschy et al., 2012; Anderson et al., 2011; Braver & Bongiolatti, 2002).

Conversely, better solving accuracy in stressful situations was associated with longer durations in a state that consisted of reduced global gamma power across all sources. This is consistent with past research indicating that widespread reductions in global activation, particularly in lateral pFC and intraparietal cortex, are evident during problem solving in stressful contexts (van Ast et al., 2016; Young, Wu, & Menon, 2012), but how this relates to better performance is less clear. One possibility is that, in stressful situations, increased activity in brain regions associated with problem solving is typically accompanied by, if not overtly competing with, activation in regions not conducive to problem solving such as

those integral for emotion regulation and self-monitoring. Hence, successful problem solving could be associated with successful attempts at stabilizing activity across the brain or “keeping the brain calm” in the face of stress. This conjecture is supported by the finding that fewer states overall were associated with success on the problem-solving task under stress.

Although the two unique brain states associated with better performance in the stressed and nonstressed conditions varied in power distribution, global synchrony analyses revealed that, compared with any other state in each condition, the two specific states were more strongly correlated with all other sources; that is, there were higher global interactions within each state. This is consistent with past research suggesting that global connectivity, especially connectivity between frontoparietal regions, increases during various cognitive tasks (Vakhtin et al., 2014; Rosenberg-Lee et al., 2011; Sauseng et al., 2005). Our results add to this literature by demonstrating that longer durations in these two states, and better global synchrony, are ultimately beneficial for performance in neutral or stressful contexts, respectively.

Findings from these studies also prompt the question: What exactly do higher numbers of HMM states represent neurally with respect to performance? Given that individuals exhibited more states on hard problems compared with easy problems, stressed individuals exhibited higher numbers of states on incorrectly solved problems, and all individuals performed worse to the extent that more states were exhibited—it seems likely that larger numbers of states represent global instability. It is less clear, however, whether this instability is the product of a maladaptive neural response or simply a dynamic system that attempts to engage multiple, possibly competing subnetworks otherwise integral for performance on cognitively intensive tasks. It is possible that state counts reflect both of these processes, but yoking them to performance in specific contexts provides insight into whether changes in global network activity are beneficial for performance. Past research suggests that stressful, evaluative contexts like the one primed in this study elicit a cascade of stress responses, performance monitoring, and appraisal processes that tax working memory resources otherwise needed for optimal performance on more difficult tasks (Schmader et al., 2008). Responses like these would likely engender the recruitment of multiple neural networks, including those integral for emotion regulation and working memory, in addition to those integral for problem solving in general. This competition would undoubtedly have ramifications for performance, but future research would be necessary to provide direct evidence for this conjecture.

It is worth noting that, given that many different cognitive and emotional processes may be active during problem solving under stress, there are of course alternative interpretations of these findings. One plausible explanation is that mind wandering or distraction played a

role in poorer performance among individuals under stress, which indeed has been shown to play a role in underperformance in stereotype threatening contexts (Mrazek et al., 2011). One would expect fleeting thoughts to be associated with an increased number of brain states in general (as well as suboptimal performance), although, to our knowledge, no direct evidence of this exists. Unfortunately, such possibilities represent the possible pitfalls of both HMM studies and cognitive neuroscience in general in that they ultimately rely on reverse inference. Our goal was to avoid this pitfall by explaining a given state as a function of what specific brain regions/sources were doing in relation to one another (with respect to fluctuations in gamma power), and global neural stability more broadly, as opposed to what psychological or cognitive state activity in a given brain region might represent, and map those states directly onto behavior (performance). Attempts to map complex psychological processes onto complex brain states (and rule out other processes like mind wandering compared with emotion regulation) are a challenge for future research utilizing HMM. Nevertheless, the approach outlined in this study represents one possible way to address this issue by quantifying representative brain states across individuals or conditions and using these outcomes to predict specific behavioral outcomes statistically.

Another important aspect of findings is that they were specific to gamma power (although similar patterns were found in theta and alpha frequency bands, see footnote 4). Frequency rhythm has been widely linked to human cognitions (Liu, Kuo, & Chiu, 2013). We focused on gamma oscillations because they appear to play an integral role in information processing and synthesizing information throughout the brain and more conscious, deliberative processes in general (Buzsáki, 2006). This is consistent with past research indicating that gamma waves are highly associated with various problem-solving processes (Jung-Beeman et al., 2004; Gruber, Muller, & Keil, 2002). This is not to say that other frequency bands are not important, however, particularly given that evidence from fMRI studies indicate that low-frequency bands within and between brain regions play an important role in the problem-solving process (Anderson et al., 2014; Krueger et al., 2008; Jung-Beeman et al., 2004). Future research is obviously needed to test more on these predictions with respect to global network stability and the problem-solving process.

Regarding limitations of the study, as with any EEG study that utilizes a source localization approach, it is always important to stress caution with respect to conjectures based on specific regions of the brain given limitations in spatial localization associated with the methodology. Nevertheless, given our interest in the timing of gamma power oscillations in relation to multiple sources across the brain, EEG represents a clear choice for our analyses. Furthermore, all sources were restricted to the outer cortex, which allows for fairly precise measurements of specific regions (Cohen, 2014). Future re-

search should replicate this study utilizing combined EEG–fMRI methodologies to allow for both optimal temporal and spatial resolution to bolster claims accordingly.

In this study, HMM-based MVPA was utilized as a novel means for measuring global network stability and subtle, transient changes across the brain during the problem-solving process. Findings suggest that more complex cognitive tasks elicit a global dynamic alteration in neural oscillatory activity, particularly in stressful contexts, that is impossible to detect with other methodologies like psychophysiological interaction analyses. Although successful problem solving was associated with an optimal state count regardless of context (but via different neural patterns), individuals under stress underperformed to the extent that they exhibited global network instability (higher state counts, fluctuations in gamma power across sources). Nonstressed individuals underperformed to the extent that they exhibited too few states and decreases in gamma power in regions integral for problem solving and executive function. Findings highlight how HMM–MVPA can provide myriad possibilities for quantifying and better understanding the role of global network interactions in the problem-solving process and how the said interactions predict success or failure in different contexts.

Reprint requests should be sent to Mengting Liu, Department of Psychological and Brain Sciences, University of Delaware, 108 Wolf Hall, Newark, DE 19716, or via e-mail: mliu@psych.udel.edu.

Notes

1. We define stress here as the feeling that situational demands exceed perceived resources. We chose to use the term “stress” as opposed to, say, anxiety because stress represents the pithiest and most direct and accurate expression of what is likely occurring in our stereotype threat condition. Although it is possible that feelings of stress and anxiety were evident in our stereotype threat condition given that evidence for both have been found in previous stereotype threat studies (e.g., Schmader et al., 2009; Johns, Inzlicht, & Schmader, 2008; Osborne, 2006, 2007; Bosson, Haymovitz, & Pintel, 2004; Blascovich, Spencer, Quinn, & Steele, 2001; Spencer et al., 1999), there is reason to believe that these psychometric terms are difficult to discern between neurally and in fact yield fairly comparable neural patterns regardless. To demonstrate this, we conducted a direct effect size-wise meta-analytic contrast between the terms “stress” and “anxiety” using Neurosynth. No significant contrasts were found in any voxels across the whole brain, indicating that stress and anxiety were highly correlated at the neural level or at least with respect to whatever way these terms were described psychometrically in the studies included in the meta-analysis. Thus, for the purposes of this study, we utilize the term “stress” to describe our stereotype threat condition and findings, but it would be difficult, if not impossible, to rule out a role for anxiety as well (at least with respect to the data available in this study).
2. Power in vACC was significantly higher in the stress condition compared with the control condition within the theta and beta frequency bands as well ($ps < .035$).
3. One participant was excluded from this analysis after being identified as an outlier in state count (collapsed across the four

conditions) via Grubbs test. When this participant was included in analyses, patterns were consistent but marginal, $p = .10$.

4. Quadratic regression analyses were also conducted on other frequency bands. Marginal significant quadratic effects were found in theta ($p = .078$) and alpha ($p = .101$) bands, but no significant effect was found in beta band ($p = .543$).

REFERENCES

- Anderson, J. R., Betts, S., Ferris, J. L., & Fincham, J. M. (2011). Cognitive and metacognitive activity in mathematical problem solving: Prefrontal and parietal patterns. *Cognitive, Affective & Behavioral Neuroscience, 11*, 52–67.
- Anderson, J. R., Fincham, J. M., Schneider, D. W., & Yang, J. (2012). Using brain imaging to track problem solving in a complex state space. *Neuroimage, 60*, 633–643.
- Anderson, J. R., Lee, H. S., & Fincham, J. M. (2014). Discovering the structure of mathematical problem solving. *Neuroimage, 97*, 163–177.
- Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science, 11*, 181–185.
- Ashcraft, M. H., & Faust, M. W. (1994). Mathematics anxiety and mental arithmetic performance—An exploratory investigation. *Cognition & Emotion, 8*, 97–125.
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*, 224–237.
- Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods, 7*, 127–150.
- Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science, 17*, 339–343.
- Beilock, S. L., & Gray, R. (2012). From attentional control to attentional spillover: A skill-level investigation of attention, movement, and performance outcomes. *Human Movement Science, 31*, 1473–1499.
- Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology, 41*, 174–181.
- Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science, 12*, 225–229.
- Bola, M., & Borchardt, V. (2016). Cognitive processing involves dynamic reorganization of the whole-brain network's functional community structure. *Journal of Neuroscience, 36*, 3633–3635.
- Bola, M., & Sabel, B. A. (2015). Dynamic reorganization of brain functional networks during cognition. *Neuroimage, 114*, 398–413.
- Borst, J. P., & Anderson, J. R. (2015). The discovery of processing stages: Analyzing EEG data with hidden semi-Markov models. *Neuroimage, 108*, 60–73.
- Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology, 40*, 247–255.
- Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage, 15*, 523–536.
- Buzsáki, G. (2006). *Rhythms of the brain*. Oxford, UK: Oxford University Press.
- Celeux, G. (2007). Mixture models for classification. In *Advances in data analysis* (pp. 3–14). Berlin: Springer.
- Chiu, A. W. L., Derchansky, M., Cotic, M., Carlen, P. L., Turner, S. O., & Bardakjian, B. L. (2011). Wavelet-based Gaussian-mixture hidden Markov model for the detection of multistage seizure dynamics: A proof-of-concept study. *Biomedical Engineering Online, 10*, 29.
- Cohen, M. X. (2014). Preface. In *Analyzing neural time series: Data theory and practice* (pp. xvii–xviii). Cambridge, MA: MIT Press.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31*, 968–980.
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*, 336–353.
- Fitzmaurice, G. M., Laird, N. M., & Rotnitzky, A. G. (1993). Regression models for discrete longitudinal responses. *Statistical Science, 8*, 284–299.
- Forbes, C. E., Duran, K. A., Leitner, J. B., & Magerman, A. (2015). Stereotype threatening contexts enhance encoding of negative feedback to engender underperformance and anxiety. *Social Cognition, 33*, 605–625.
- Forbes, C. E., & Leitner, J. B. (2014). Stereotype threat engenders neural attentional bias toward negative feedback to undermine performance. *Biological Psychology, 102*, 98–107.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*, 267.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2014). MNE software for processing MEG and EEG data. *Neuroimage, 86*, 446–460.
- Gruber, T., Muller, M. M., & Keil, A. (2002). Modulation of induced gamma band responses in a perceptual learning task in the human EEG. *Journal of Cognitive Neuroscience, 14*, 732–744.
- Henckens, M. J., van Wingen, G. A., Joels, M., & Fernandez, G. (2012). Time-dependent effects of cortisol on selective attention and emotional interference: A functional MRI study. *Frontiers in Integrative Neuroscience, 6*, 66.
- Hopko, D. R., McNeil, D. W., Gleason, P. J., & Rabalais, A. E. (2002). The emotional Stroop paradigm: Performance as a function of stimulus properties and self-reported mathematics anxiety. *Cognitive Therapy and Research, 26*, 157–166.
- Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General, 137*, 691.
- Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., et al. (2004). Neural activity when people solve verbal problems with insight. *PLoS Biology, 2*, 500–510.
- Krueger, F., Spampinato, M. V., Pardini, M., Pajevic, S., Wood, J. N., Weiss, G. H., et al. (2008). Integral calculus problem solving: An fMRI investigation. *NeuroReport, 19*, 1095–1099.
- Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13–22.
- Lieberman, M. D., Burns, S. M., Torre, J. B., & Eisenberger, N. I. (2016). Reply to Wager et al.: Pain and the dACC: The importance of hit rate-adjusted effects and posterior probabilities with fair priors. *Proceedings of the National Academy of Sciences, U.S.A., 113*, E2476–E2479.
- Liston, C., McEwen, B. S., & Casey, B. J. (2009). Psychosocial stress reversibly disrupts prefrontal processing and attentional control. *Proceedings of the National Academy of Sciences, U.S.A., 106*, 912–917.
- Liu, M., Kuo, C. C., & Chiu, A. W. (2011). Statistical threshold for nonlinear Granger causality in motor intention analysis. In *Engineering in Medicine and Biology Society, EMBC, 2011*

- Annual International Conference of the IEEE* (pp. 5036–5039). Boston, MA: IEEE.
- Liu, M. T., Kuo, C. C., & Chiu, A. W. (2013). Non-linear Granger causality and its frequency decomposition in decoding human upper limb movement intentions. *International Journal of Biomedical Engineering and Technology*, *12*, 1–25.
- Mrazek, M. D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology*, *47*, 1243–1248.
- Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. *Journal of Neuroscience Methods*, *192*, 152–162.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*, 424–430.
- Osborne, J. W. (2006). Gender, stereotype threat, and anxiety: Psychophysiological and cognitive evidence. *Electronic Journal of Research in Educational Psychology*, *4*, 109–138.
- Osborne, J. W. (2007). Linking stereotype threat and anxiety. *Educational Psychology*, *27*, 135–154.
- Paynter, C. A., Reeder, L. M., & Kieffaber, P. D. (2009). Knowing we know before we know: ERP correlates of initial feeling-of-knowing. *Neuropsychologia*, *47*, 796–803.
- Rabiner, L. R. (1989). A tutorial on hidden Markov-models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*, 257–286.
- Rosenberg-Lee, M., Barth, M., & Menon, V. (2011). What difference does a year of schooling make? Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving. *Neuroimage*, *57*, 796–808.
- Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., et al. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *Neuroimage*, *60*, 830–846.
- Rzucidlo, J. K., Roseman, P. L., Laurienti, P. J., & Dagenbach, D. (2013). Stability of whole brain and regional network topology within and between resting and cognitive states. *PLoS One*, *8*, e70275.
- Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, *57*, 97–103.
- Schmader, T., Forbes, C. E., Zhang, S., & Mendes, W. B. (2009). A metacognitive perspective on the cognitive deficits experienced in intellectually threatening environments. *Personality and Social Psychology Bulletin*, *35*, 584–596.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336–356.
- Simonsohn, U. (2016). Two-lines: The first valid test of U-shaped relationships. Available at http://urisohn.com/papers/u_shape.pdf.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4–28.
- Sporns, O. (2013). Structure and function of complex brain networks. *Dialogues in Clinical Neuroscience*, *15*, 247–262.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.
- Suarez-Pellicioni, M., Nunez-Pena, M. I., & Colome, A. (2013). Abnormal error monitoring in math-anxious individuals: Evidence from error-related brain potentials. *PLoS One*, *8*, e81143.
- Suarez-Pellicioni, M., Nunez-Pena, M. I., & Colome, A. (2014). Reactive recruitment of attentional control in math anxiety: An ERP study of numeric conflict monitoring and adaptation. *PLoS One*, *9*, e99579.
- Suk, H. I., Wee, C. Y., Lee, S. W., & Shen, D. (2016). State-space model with deep learning for functional dynamics estimation in resting-state fMRI. *Neuroimage*, *129*, 292–307.
- Takahashi, T., Ikeda, K., Ishikawa, M., Kitamura, N., Tsukasaki, T., Nakama, D., et al. (2005). Anxiety, reactivity, and social stress-induced cortisol elevation in humans. *Neuroendocrinology Letters*, *26*, 351–354.
- Telesford, Q. K., Morgan, A. R., Hayasaka, S., Simpson, S. L., Barret, W., Kraft, R. A., et al. (2010). Reproducibility of graph metrics in fMRI networks. *Frontiers in Neuroinformatics*, *4*, 117.
- Vakhtin, A. A., Ryman, S. G., Flores, R. A., & Jung, R. E. (2014). Functional brain networks contributing to the Parieto-Frontal Integration Theory of Intelligence. *Neuroimage*, *103*, 349–354.
- van Ast, V. A., Spicer, J., Smith, E. E., Schmer-Galunder, S., Liberzon, I., Abelson, J. L., et al. (2016). Brain mechanisms of social threat effects on working memory. *Cerebral Cortex*, *26*, 544–556.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*, 665–670.
- Young, C. B., Wu, S. S., & Menon, V. (2012). The neurodevelopmental basis of math anxiety. *Psychological Science*, *23*, 492–501.
- Yu, S. Z., & Kobayashi, H. (2003). An efficient forward-backward algorithm for an explicit-duration hidden Markov model. *IEEE Signal Processing Letters*, *10*, 11–14.
- Yu, S. Z., & Kobayashi, H. (2006). Practical implementation of an efficient forward-backward algorithm for an explicit-duration hidden Markov model. *IEEE Transactions on Signal Processing*, *54*, 1947–1951.