PSYCHOLOGICAL SCIENCE

Research Article

The Negative Consequences of Threat

A Functional Magnetic Resonance Imaging Investigation of the Neural Mechanisms Underlying Women's Underperformance in Math

Anne C. Krendl,[1] Jennifer A. Richeson,[2] William M. Kelley,[1] and Todd F. Heatherton[1]

[1]Department of Psychological and Brain Sciences, Dartmouth College, and [2]Department of Psychology and Institute for Policy Research, Northwestern University

ABSTRACT—***This study used functional magnetic resonance imaging to identify the neural structures associated with women's underperformance on math tasks. Although women in a control condition recruited neural networks that are associated with mathematical learning (i.e., angular gyrus, left parietal and prefrontal cortex), women who were reminded of gender stereotypes about math ability did not recruit these regions, and instead revealed heightened activation in a neural region associated with social and emotional processing (ventral anterior cingulate cortex).***

The underlying cause of the disparity in the number of men and women in math and science careers has been widely and vigorously debated in recent years. Potential explanations have ranged from suppositions of innate sex differences that predispose men more than women to succeed in math and science, to claims of systematic biases in the educational system. In the present article, we consider another possible explanation: the effect of stereotypes about putative sex differences in math ability.

In an influential study, Steele and Aronson (1995) found that reminding people about a negative stereotype of their social group in a specific domain (e.g., African Americans and intellectual ability) impaired their subsequent performance on related tasks, an effect Steele and Aronson referred to as stereotype threat. Likewise, Spencer, Steele, and Quinn (1999) observed that women's performance on a difficult math test suffered after they were told the test had previously revealed sex differences in performance.

A large body of research has subsequently emerged to explore the effects of stereotype threat on women's math performance. Although this research has uncovered a wealth of information about various methods of inducing stereotype threat (Davies, Spencer, Quinn, & Gerhardstein, 2002; Inzlicht & Ben-Zeev, 2000; O'Brien & Crandall, 2003; Sekaquaptewa & Mischa, 2003) and potential ways to minimize the effects of the threat (Good, Aronson, & Inzlicht, 2003; Johns, Schmader, & Martens, 2005), the mechanisms underlying stereotype threat are not fully understood (but see Wraga, Helt, Jacobs, & Sullivan, 2007).

Recent behavioral research has implicated several possible mechanisms—most notably, working memory and anxiety—that might be responsible for the performance decrements associated with stereotype threat (Beilock, Jellison, Rydell, McConnell, & Carr, 2006; Osborne, 2001; Pronin, Steele, & Ross, 2004; Schmader, 2002; Schmader & Johns, 2003). Schmader and Johns (2003) found that women under stereotype threat demonstrated lower working memory capacity and poorer performance on a difficult math test, compared with control subjects. Cadinu, Maass, Rosabianca, and Kiesner (2005) recently demonstrated that women under stereotype threat experience more domain-specific negative thoughts when performing difficult math tests, compared with control subjects. Cadinu et al. argued that such negative thoughts consume working memory resources. Inzlicht and Ben-Zeev (2003), however, argued that stereotype threat increases anxiety, which subsequently impairs performance. They found that women underperformed on math

---

Address correspondence to Anne C. Krendl, 6207 Moore Hall, Dartmouth College, Hanover, NH 03755, e-mail: akrendl@dartmouth.edu.

tests when they were the sole female in a group of males. Similarly, Sekaquaptewa and Mischa (2003) demonstrated that women who believed they would have to take an oral exam in front of a room of men performed worse on the test than women who believed they would have to take the exam in front of a room of other women. The authors suggested that this finding is further evidence that stereotype threat induces performance anxiety.

Together, these findings suggest a role for both working memory and performance anxiety in stereotype threat. However, it remains unclear whether their roles are interconnected. Beilock, Rydell, and McConnell (2007) suggested that working memory and performance anxiety are activated separately. Specifically, they posited that stereotype threat causes performance anxiety, and that anxiety subsequently triggers negative math-relevant thoughts that consume working memory. In other words, working memory deficits accompany stereotype threat because the threat heightens awareness of and concerns about the implications of the stereotypes. Support for this argument came from an earlier study (reported by Beilock et al., 2006) showing that performance on tasks that do not rely on working memory is negatively affected by stereotype threat.

Although the findings by Beilock et al. (2007) and other researchers have greatly informed the stereotype-threat literature, it remains unclear how the mechanisms that are believed to underlie stereotype threat cause subsequent performance decrements. In order to effectively override the stereotype-threat phenomenon, it is vital to understand clearly the core processes that underlie it. In pursuing this goal, researchers may benefit from using neuroimaging to identify brain regions engaged during stereotype threat. When considered in tandem with the extant behavioral research, neuroimaging studies may help elucidate the neural mechanisms engaged during stereotype threat and, furthermore, dissociate them from the neural mechanisms engaged during performance of relevant tasks in the absence of stereotype threat. Thus, in the current study, we used functional magnetic resonance imaging (fMRI) to identify the neural processes engaged when women perform difficult math tasks both in the presence and in the absence of stereotype threat.

METHOD

Participants

The participants were recruited for a study measuring “the neural mechanisms engaged in cognitive tasks that require both speed and accuracy.” All participants were right-handed female undergraduates at Dartmouth College (*N* = 28; 14 control subjects) and were highly identified with math (Aronson et al., 1999), as determined by their response to the following question: “It is important to me that I am good at math.” A response of 4 or higher on a 7-point Likert scale was required for participation. Upon arriving at the Dartmouth Brain Imaging Center, participants were randomly assigned to the control or threat condition. They received partial course credit or $20 remuneration for participating. A male experimenter who was blind to condition recruited the participants, gave them generic instructions for the task, and put them in the scanner. During the experiment, specific instructions that varied between conditions were presented by computer.

Behavioral Procedure

When participants arrived at the lab, the male experimenter instructed them that they would complete a series of categorization tasks and evaluate math equations. They were told that each math problem would be presented as a solved equation (e.g., “Is $5 \times 2 - 3 = 7$?”), and that they would be given 5 s to determine if the equation was true or false. They were to indicate their decision via key press. The math problems were pseudorandomly intermixed with trials on which participants were required to look at a jittered fixation point. The latter trials established a baseline that allowed for subsequent event-related analysis.

The session began with a neutral version of the Implicit Association Test (IAT), a widely used implicit measure of attitudes (Greenwald, McGhee, & Schwartz, 1998). The IAT requires participants to categorize words as belonging to one of four categories (in this case, flowers, insects, pleasant, or unpleasant). The categories are presented in congruent blocks (i.e., “flowers” and “pleasant” paired together) and incongruent blocks (i.e., “insects” and “pleasant” paired together). Categories that were paired together were presented on the same side of the screen, and participants were required to use the same response key to categorize stimuli belonging to these categories (e.g., when “flowers” and “pleasant” were paired, participants pressed the same key to categorize the word *tulip* as a flower and to categorize the word *rainbow* as pleasant); implicit bias was measured by calculating the difference in response times between the congruent and incongruent blocks.

Upon completion of the IAT, participants were given an initial set of 50 difficult math problems that involved basic arithmetic: addition, subtraction, multiplication, division, and square exponents (e.g., “Is $19 \times 6 - 6 \wedge 2 = 78$?” and “Is $98/7 + 19 \times 3 = 81$?”). We also included modular arithmetic problems within the set of 50 problems (Gauss, 1966, p. 472). In these problems, the task was to subtract the second number from the first and divide the result by the third number, in parentheses. For example, to solve the problem $37 = 18 \pmod 6$, participants would subtract 18 from 37 and then divide by 6. Participants were told to indicate via key press whether or not the solution to each modular problem was a whole number. Pilot testing revealed that presenting the difficult math problems intermixed with modular problems yielded 56.7% accuracy within the 5-s time constraint.

After finishing the first set of math problems, participants were told they would complete another categorization task. In the threat condition, participants were told they were going to complete a task that would assess their "math attitudes" because "research has shown gender differences in math ability and performance." These instructions served as the primary threat induction, as previous research has suggested that reminding women of gender stereotypes in math ability activates stereotype threat (e.g., O'Brien & Crandall, 2003; Spencer et al., 1999, Study 2). However, because this study was the first attempt to induce stereotype threat while participants were in the scanner, the second categorization task we administered in this condition was the math/arts IAT (Nosek, Banaji, & Greenwald, 2002b), in which participants categorized words as being related to math, arts, males, or females. We intended for this task to further reinforce the salience of gender stereotypes regarding math ability, and indeed, pilot testing revealed that administering this IAT was sufficient to induce the stereotype-threat effect.

In the control condition, participants were told they would complete a categorization task to assess their "political attitudes" to control for "individual differences in how personal attitudes . . . modulate performance on cognitive tasks." Participants were then given the instructions for the liberal/conservative IAT, which involved categorizing words as being related to liberals, conservatives, unpleasant, or pleasant (Nosek, Banaji, & Greenwald, 2002a). Thus, control participants received no reminder of gender stereotypes regarding math ability.

Following the completion of the second IAT, both the threat and the control groups were given a second math test containing another 50 problems. The problems in the second test were similar to those described previously.

Imaging Parameters

All tasks were performed while participants were in the scanner. Imaging was performed on a 3.0-T Philips Intera scanner using an eight-channel phase arrayed coil. Visual stimuli were generated with an Apple G3 laptop computer running PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). They were projected onto a screen positioned at the head end of the bore by an Epson (Model ELP-7000) LCD projector. Participants viewed the screen through a mirror mounted on top of the head coil. A fiber-optic, light-sensitive key press that interfaced with the PsyScope button box (New Micros, Dallas, TX) was used to record participants' behavioral responses. Cushions minimized head movement.

Anatomical images were acquired using a high-resolution 3-D spoiled gradient recovery sequence (160 sagittal slices, repetition time = 9.9 ms, echo time = 4.6 ms, flip angle = 8°, voxel size = 0.94 × 0.94 × 1 mm). Functional images were collected using a gradient spin-echo-planar sequence sensitive to blood-oxygenation-level-dependent (BOLD) contrast (T_2^*; repetition time = 2,500 ms, echo time = 35 ms, flip angle = 90°, 3 × 3 mm in-plane resolution). Data were collected in one functional run with approximately 500 sets of axial images (20 slices; slice thickness of 3.5 mm, 0.5-mm skip between slices).

Imaging Data Analysis

The fMRI data were analyzed using Statistical Parametric Mapping software (Friston et al., 1995). Data collected during the functional run were preprocessed to remove sources of noise and artifact. Functional data were corrected for differences in acquisition time between slices for each whole-brain volume, realigned within and across runs to correct for head movement, and coregistered with each participant's anatomical data. Functional data were then transformed into a standard anatomical space (2-mm isotropic voxels) based on the International Consortium of Brain Mapping 152 brain template (Montreal Neurological Institute), which approximates Talairach and Tournoux (1988) atlas space. Normalized data were then spatially smoothed (6-mm full width at half maximum) using a Gaussian kernel.

For each participant, a general linear model incorporating task effects (Friston et al., 1998) and covariates of no interest (a mean for the functional run, a linear trend for the functional run, and six movement parameters derived from realignment) was used to compute parameter estimates (β) and t-contrast images (containing weighted parameter estimates) for each comparison at each voxel. The two sets of math problems were analyzed as events, and the two IATs were analyzed as blocks. The individual contrast images were then submitted to a second-level, random-effects analysis to create mean t images (threshold at $p < .001$, uncorrected). An automated search algorithm identified the location of peak activations and deactivations on the basis of z values and cluster sizes.

To analyze the math event data, we used each individual's design matrix to compute parameter estimates (β) and t-contrast images (containing weighted parameter estimates) for each comparison at each voxel. The main contrasts of interest compared brain activation during the second math test with activation during the first math test. These contrast images were used for subsequent region-of-interest (ROI) analyses.

ROI analyses were conducted using the functional ROIs tool in SPM99 (Wellcome Department of Imaging Neuroscience, London, United Kingdom). All significant voxels ($p < .001$) within 6 mm of a peak location were included in each ROI. An extent threshold of five contiguously activated voxels was also applied. Signal intensities for each ROI were then calculated separately for each condition and examined statistically using repeated measures analysis of variance (ANOVA).

RESULTS

Behavioral Results

Overall Performance

Performance (total number of math items answered correctly) was assessed in an ANOVA with time (Time 1 or Time 2) as a within-subjects variable and condition (control or threat) as a between-subjects variable. The ANOVA revealed a significant interaction, $F(1, 26) = 11.41, p < .005, \eta_p^2 = .31$ (Fig. 1), and

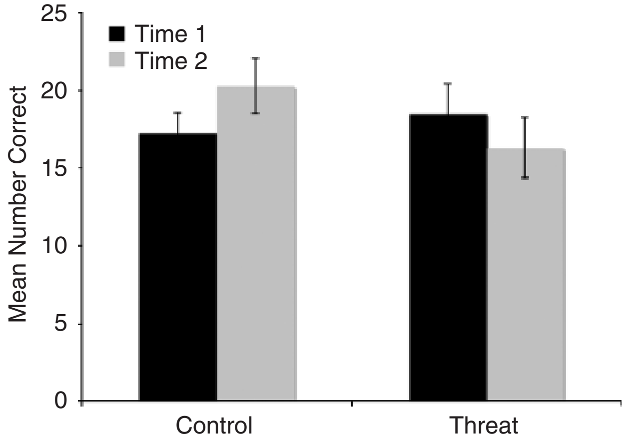

Fig. 1. **Accuracy of control and threatened participants as a function of time (Time 1 and Time 2). For each participant, performance was measured as the number of problems correct.**

no main effect of time ($F < 1$) or condition, $F(1, 26) = 1.24, p = .28$. Subsequent t tests indicated that the interaction emerged because the performance of control participants improved significantly over time, $t(13) = 2.81, p < .02$, whereas the threatened group's performance decreased slightly over time, $t(13) = 1.98, p = .07$.[1]

Reaction Times

We used a standard log transformation to normalize the reaction times, which are inherently otherwise positively skewed. The mean log reaction times for each participant were then analyzed in a mixed-model ANOVA patterned after the analysis of overall performance. This ANOVA revealed a main effect of time, $F(1, 26) = 6.21, p < .02, \eta_p^2 = .19$, but no effect of condition ($F < 1$), and no interaction, $F(1, 26) = 1.21, p = .28$. Subsequent t tests revealed that control participants' reaction times were significantly faster on the second math test than on the first, $t(13) = 3.91, p < .01$. Participants in the threatened group, however, showed no significant difference in reaction time between the first and second tests. Reaction time was not associated with accuracy.

fMRI Results

From the imaging data, we identified neural activations that were greater during the second math test than the first (see Tables 1 and 2). These analyses showed that the control participants recruited more left-lateralized activation in the inferior prefrontal cortex (Brodmann's area, BA, 47), left inferior parietal cortex (BA 40), and bilateral angular gyrus (BA 39) over time. By contrast, the threatened participants revealed greater activity in the ventral anterior cingulate cortex (vACC; BA 32/10) on the second test than on the first (Fig. 2).

In order to determine if there were functional dissociations between these ROIs, we conducted a 2 (condition: control or threat) × 2 (time: Time 1 or Time 2) ANOVA for each ROI identified in the preceding paragraph. Significant interactions emerged for BA 47, $F(1, 26) = 7.92$; left BA 39, $F(1, 26) = 8.96$; and right BA 39, $F(1, 26) = 7.35$, all $ps < .02$; there was also a trend for BA 40, $F(1, 26) = 2.93, p < .1$. Post hoc statistical tests revealed that the two-way interactions emerged because control participants recruited BA 47, BA 40, and BA 39 to a greater extent over time ($ps < .001$), whereas threatened participants did not. A separate ANOVA conducted on vACC activation revealed a significant interaction, $F(1, 26) = 5.97$, $p = .02$; threatened participants, but not control participants, recruited heightened vACC activation over time, $t(13) = 5.64, p < .001$.

These findings suggest a functional dissociation between vACC and the other ROIs. To formally test for this functional dissociation, we computed additional ANOVAs that included region as a third factor. Specifically, we conducted 2 (time: Time 1 or Time 2) × 2 (condition: control or threat) × 2 (region: vACC vs. each cognitive region) mixed-model ANOVAs with time and region as within-subjects variables and condition as a between-subjects variable. In these analyses, functional dissociations between regions could occur because (a) the regions behaved differently across conditions (resulting in a Region × Condition interaction), (b) the regions behaved differently across time (causing a Region × Time interaction), or (c) the regions demonstrated different Condition × Time interactions (resulting in a Region × Condition × Time interaction). We found a significant three-way interaction in each ANOVA—BA 47: $F(1, 26) = 13.94$; left BA 40: $F(1, 26) = 10.99$; left BA 39: $F(1, 26) = 11.31$; right BA 39: $F(1, 26) = 13.39$ (all $ps < .005$; see Figs. 2c and 2d). The three-way interactions show that regions exhibiting a significant change in activation over time in one group of participants did not do so in the other group of participants.

Together, these interactions revealed a double dissociation between cognitive regions (left inferior frontal, left parietal, and bilateral angular gyrus regions) and an affective region (vACC). Whereas left inferior frontal, left parietal, and bilateral angular gyrus regions demonstrated increased activity for control participants but not threatened participants over time, vACC demonstrated the reverse pattern (i.e., increased activity for threatened but not control participants over time).

DISCUSSION

The results obtained are consistent with predictions from the stereotype-threat literature. Women who were reminded of the stereotype that there are sex differences in math ability underperformed on subsequent math problems compared with women who received no such reminder. These behavioral differences were accompanied by robust differences in neural activity between the two conditions.

---

[1] Because performance differences emerged at Time 1, we conducted an analysis of covariance with Time 1 performance as a covariate. A significant effect of condition still emerged, demonstrating that the interaction was not driven by group differences at Time 1.

TABLE 1

Regions of Interest Exhibiting Greater Activation at Time 2 Than at Time 1 Among Control Participants

Brain region	Coordinates x	Coordinates y	Coordinates z	t score	No. of voxels
Left inferior parietal gyrus (BA 40)	−53	−56	44	7.54	17
Left angular gyrus (BA 39)	−53	−72	15	5.99	11
Right angular gyrus (BA 39)	53	−69	28	4.39	5
Left inferior frontal gyrus (BA 47)	−45	35	−2	5.66	5
Left medial frontal gyrus (BA 6/8)	−36	14	44	6.29	10
Right precentral gyrus (BA 4)	27	−26	70	6.26	10
Left superior temporal gyrus (BA 22)	−65	−23	1	5.28	7
Left cerebellum	−21	−45	−33	4.9	5
Pons	−12	−28	−31	6.2	5
Left caudate	−12	15	2	4.65	5

Note. To identify these regions, we used a significance threshold of $p < .001$, uncorrected, and a 5-voxel extent threshold. BA = Brodmann's area.

Among women in the control condition, the brain regions that revealed heightened activation over time are consistent with the regions that previous imaging studies have identified as being associated with math computations, such as mathematical calculations (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999), spatial representation of numbers (Gobel, Walsh, & Rushworth, 2001), and mental rotation (Halari et al., 2006). These regions have been implicated in math learning. Delazer et al. (2003) found that the angular gyrus (specifically, on the left) was active in participants who were highly trained on difficult math tests, a result indicating that the angular gyrus is critical to developing mathematical learning skills. Menon, Rivera, White, Glover, and Reiss (2000) found heightened activation in the angular gyrus as the difficulty of the math task increased. Interestingly, they also observed heightened activation in caudate, thalamus, and cerebellum for increasingly difficult tasks, a pattern of activation that we also observed in our control participants (see Table 1).

The predominantly left lateralization of neural activation over time in our control participants is also consistent with previous research suggesting that right-handed participants rely on left prefrontal areas (i.e., inferior and medial frontal gyri) to perform mental math tasks (Burbaud et al., 1995) and left lateralized parietal cortex when they have begun to develop an expertise on specific math tasks (Delazer et al., 2003). Together, these findings provide compelling evidence that participants in our control condition were developing strategies that enhanced their performance on the math task, as evidenced by their behavioral improvement over time.

TABLE 2

Regions of Interest Exhibiting Greater Activation at Time 2 Than at Time 1 Among Threatened Participants

Brain region	Coordinates x	Coordinates y	Coordinates z	t score	No. of voxels
Ventral anterior cingulate (BA 32/10)	−6	37	−22	6.26	11
Orbitofrontal cortex (BA 11)	30	40	−17	5.63	6
Right postcentral gyrus (BA 2)	42	−32	65	5.19	7

Note. To identify these regions, we used a significance threshold of $p < .001$, uncorrected, and a 5-voxel extent threshold. BA = Brodmann's area.

The women who received the stereotype-threat reminder, however, failed to show increased recruitment of these mathematical brain regions and also performed worse over time. Unlike the control condition, the threat condition was associated with robust activity in the vACC. Previous research has implicated this region as an affective network that is selectively engaged in assessing the salience of emotional information and initiating emotion-regulation processes (Bush, Luu, & Posner, 2000). In a recent study, Somerville, Heatherton, and Kelley (2006) demonstrated that the vACC is uniquely sensitive to social feedback, such as social rejection. Moran, Macrae, Heatherton, Wyland, and Kelley (2006) found that the vACC selectively responded to evaluating negative self-relevant trait words during a social-processing task. Interestingly, the vACC has also been extensively implicated in clinical depression. Drevets et al. (1997) found that clinically depressed patients had an abnormally hypometabolic vACC, and Mayberg et al. (2005) discovered that stimulating the vACC appeared to improve the emotional states of several clinically depressed patients. These studies suggest that the vACC plays a central role in processing negative social information. Accordingly, it is not surprising to see brain activity in vACC increase during a social threat. Together, these findings shed light on the mental operations that underlie the pernicious effects of stereotype threat. Specifically, in the presence of stereotype threat, women expe-

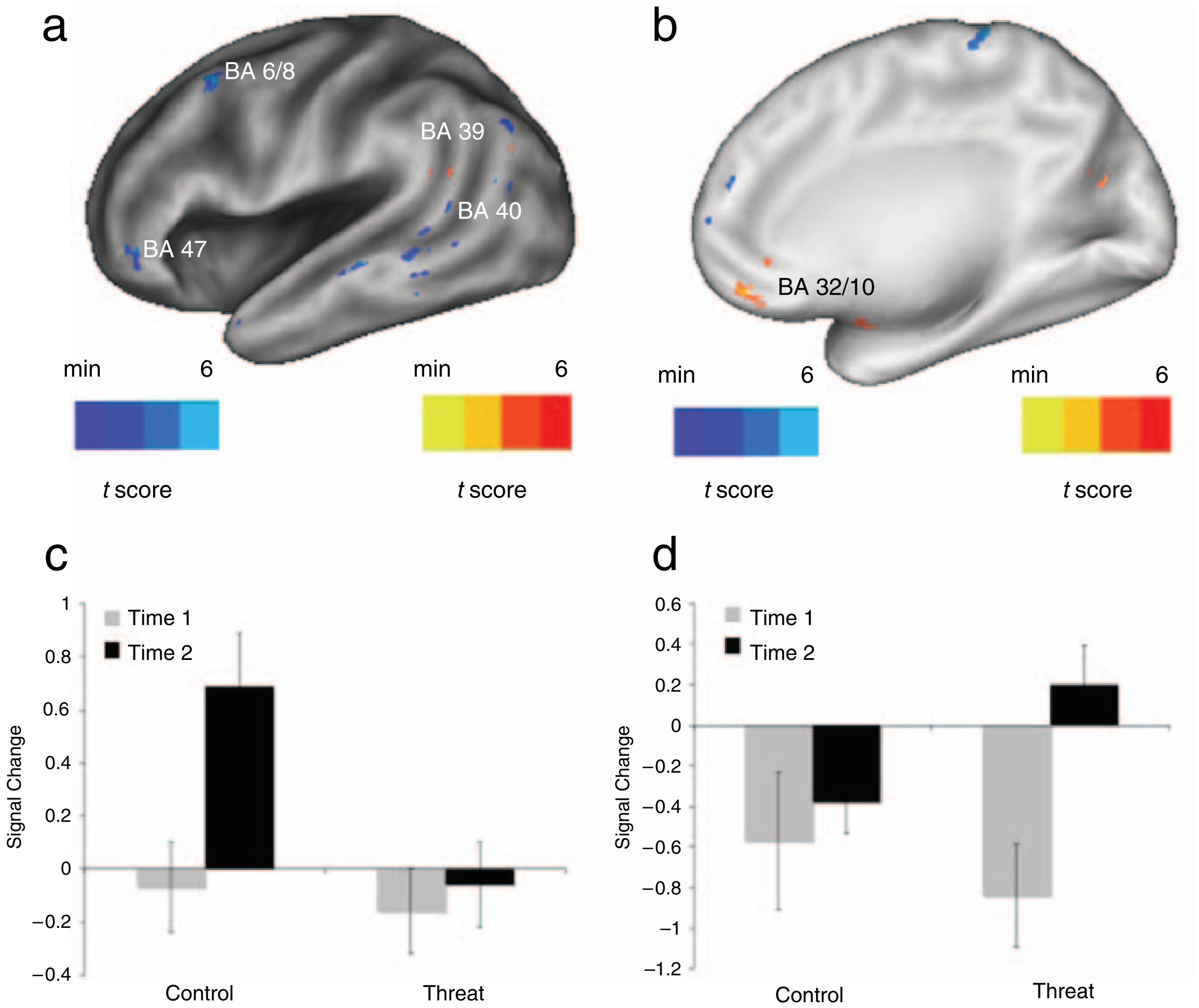

Fig. 2. Changes in neural activation over time for control and threatened participants. The lateral view of the left hemisphere of an inflated brain (a) and medial view of the right hemisphere of an inflated brain (b) depict neural regions that were more active during the second math task than the first. Activation for control participants is depicted in blue, whereas activation for threatened participants is depicted in yellow, orange, and red. Brodmann's areas (BAs) are indicated. The activation threshold was $p < .001$, uncorrected, with a minimum (''min'') t of 3.5 and a maximum t of 6.0. The bar graphs depict signal change (from a fixation control task) as a function of condition (control or threat) and time (Time 1 or Time 2) for left inferior prefrontal cortex (c; BA 47) and ventral anterior cingulate cortex (d; BA 32/10).

rience heightened activation in the vACC, a region implicated in social and emotional processing, and subsequently fail to adopt effective mathematical learning strategies.

With respect to the predominant theories about the mechanisms that give rise to stereotype-threat effects, our findings offer two important insights: First, stereotype threat does not necessarily increase demands on working memory (as evidenced by the fact that neural regions that support working memory were no more active after the threat induction than before). Second, a primary neural consequence of stereotype threat is to increase activation in the vACC, a region involved in social and emotional processing. We are certainly not suggesting that working memory is unaffected by stereotype threat. Rather, the heightened vACC activation we observed may occur at the expense of successfully recruiting other neural networks, such as regions involved in mathematical learning or working memory, as suggested by Beilock et al. (2007).

The primary effects of stereotype threat observed in this study may reflect several different psychological processes. For instance, stereotype threat may direct women's attention toward the negative social and emotional consequences of confirming negative stereotypes about their group, thereby increasing performance anxiety. Such a finding would be consistent with the abundant research documenting the effects of social anxiety on the performance of complex tasks (Zajonc, 1965). Stereotype threat may also cause women to overmonitor their performance. Being reminded of negative stereotypes about their group may heighten women's awareness of their performance and limit the cognitive resources they devote to the task (e.g., Beilock et al., 2007).

Although our findings cannot address what thoughts were associated with the vACC activation observed in the threatened participants, they advance the extant literature on stereotype threat. Extensive research on stereotype threat has been conducted across myriad domains and for a wide range of groups, and the conclusions have been very much the same: Reminding a group about negative stereotypes related to their abilities to perform a task has a deleterious impact on their ability to perform that task. This line of research has shaped understanding of the pervasive and maleficent effects of stereotype threat, but has only begun to enhance understanding of the underlying mental processes affected. The present study sheds light on the specific neural mechanisms affected by stereotype threat, thereby providing considerable promise for developing methods to forestall its unintended consequences.

Acknowledgments—The authors thank Bryan Denny for his assistance in data collection, and Emily Cross, David Kramer, and Brian Russ for their assistance with data analysis. Support for this research was provided by National Science Foundation Grant SBE-035440, the Dartmouth Brain Imaging Center, and a National Science Foundation Graduate Fellowship to A.C.K.

REFERENCES

Aronson, J., Lustina, M.J., Good, C., Keough, K., Steele, C.M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, *35*, 29–46.

Beilock, S.L., Jellison, W.A., Rydell, R.J., McConnell, A.R., & Carr, T.H. (2006). On the causal mechanisms of stereotype threat: Can skills that don't rely heavily on working memory still be threatened? *Personality and Social Psychology Bulletin*, *32*, 1059–1071.

Beilock, S.L., Rydell, R.J., & McConnell, A.R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, *136*, 256–276.

Burbaud, P., Degreze, P., Lafon, P., Franconi, J.M., Bouligand, B., Bioulac, B., et al. (1995). Lateralization of prefrontal activation during internal mental calculation. *Journal of Neurophysiology*, *74*, 2194–2200.

Bush, G., Luu, P., & Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215–222.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, *16*, 572–578.

Cohen, J.D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, *25*, 257–271.

Davies, P.G., Spencer, S.J., Quinn, D.M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, *28*, 1615–1628.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*, 970–974.

Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic—an fMRI study. *Cognitive Brain Research*, *18*, 76–88.

Drevets, W.C., Price, J.L., Simpson, J.R., Todd, R.D., Reich, T., Vannier, M., & Raichle, M.E. (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature*, *386*, 824–827.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *NeuroImage*, *7*, 30–40.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., & Frackowiak, R.S.J. (1995). Statistical parametric maps in functional imaging: A general linear model approach. *Human Brain Mapping*, *2*, 189–210.

Gauss, C.F. (1966). *Disquisitiones arithmeticae*. New Haven, CT: Yale University Press.

Gobel, S., Walsh, V., & Rushworth, M.F.S. (2001). The mental number line and the human angular gyrus. *NeuroImage*, *14*, 1278–1289.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*, 645–662.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Halari, R., Sharma, T., Hines, M., Andrew, C., Simmons, A., & Kumari, V. (2006). Comparable fMRI activity with differential behavioral performance on mental rotation and overt verbal fluency tasks in healthy men and women. *Experimental Brain Research*, *169*, 1–14.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, *11*, 365–371.

Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology*, *95*, 796–805.

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, *16*, 175–179.

Mayberg, H.S., Lozano, A.M., Voon, V., McNeely, H.E., Seminowicz, D., Hamani, C., et al. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, *45*, 651–660.

Menon, V., Rivera, S.M., White, C.D., Glover, G.H., & Reiss, A.L. (2000). Dissociating prefrontal and parietal cortex activation during arithmetic processing. *NeuroImage*, *12*, 357–365.

Moran, J.M., Macrae, C.N., Heatherton, T.F., Wyland, C.L., & Kelley, W.M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, *18*, 1586–1594.

Nosek, B.A., Banaji, M.R., & Greenwald, A.G. (2002a). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*, 101–115.

Nosek, B.A., Banaji, M.R., & Greenwald, A.G. (2002b). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44–59.

O'Brien, L.T., & Crandall, C.S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*, 782–789.

Osborne, J.W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology*, *26*, 291–310.

Pronin, E., Steele, C.M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, *40*, 152–168.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, *38*, 194–201.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*, 440–452.

Sekaquaptewa, D., & Mischa, T. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, *39*, 68–74.

Somerville, L.H., Heatherton, T.F., & Kelley, W.M. (2006). Dissociating expectancy violation from social rejection. *Nature Neuroscience*, *9*, 1007–1008.

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.

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.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain* (M. Rayport, Trans.). New York: Thieme Medical.

Wraga, M., Helt, M., Jacobs, E., & Sullivan, K. (2007). Neural basis of stereotype-induced shifts in women's mental rotation performance. *Social Cognitive and Affective Neuroscience*, *2*, 12–19.

Zajonc, R.B. (1965). Social facilitation. *Science*, *149*, 269–274.

(Received 4/17/07; Revision accepted 7/26/07)