



Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

FlashReport

Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE

Jeremy P. Jamieson^a, Wendy Berry Mendes^{a,*}, Erin Blackstock^a, Toni Schmader^b^a Department of Psychology, Harvard University, United States^b Department of Psychology, University of British Columbia, Canada

ARTICLE INFO

Article history:

Received 17 July 2009

Revised 20 August 2009

Available online 3 September 2009

Keywords:

Reappraisal

Physiological reactivity

Emotion–cognition interactions

Stress and performance

ABSTRACT

This research examined the benefits of interpreting physiological arousal as a *challenge* response on practice and actual Graduate Record Examination (GRE) scores. Participants who were preparing to take the GRE reported to the laboratory for a practice GRE study. Participants assigned to a reappraisal condition were told arousal improves performance, whereas control participants were not given this information. We collected saliva samples at baseline and after the appraisal manipulation, which were then assayed for salivary alpha amylase (sAA), a measure of sympathetic nervous system activation. Reappraisal participants exhibited a significant increase in sAA and outperformed controls on the GRE-math section. One to three months later, participants returned to the lab and provided their score reports from their actual GRE. Again, reappraisal participants scored higher than controls on the GRE-math section. These findings illuminate the powerful influence appraisal has on physiology and performance both in and out of the laboratory.

© 2009 Elsevier Inc. All rights reserved.

Introduction

“Not everything that counts can be counted and not everything that can be counted counts.” – Albert Einstein

Although high-stakes standardized tests, such as the SAT and Graduate Record Examination (GRE), influence whether students will be accepted to or rejected from desired academic programs, the above quote illustrates the necessity of considering factors other than aptitude and ability when evaluating standardized test performance. For instance, test-takers may feel an increase in arousal, or “nervous energy,” which may be interpreted as anxiety or threat, and be associated with poor performance (Cassaday & Johnson, 2002).

However, arousal is a fuzzy term semantically and psychologically (Blascovich, 1992). Arousal increases co-occur with a variety of emotional, cognitive, and motivational states and do not necessarily indicate a negative state such as anxiety or threat. Arousal increases can also indicate that the body is mobilizing resources to meet the task demands and could signal an approach orientation or *challenge* response. Because of its association with both benign and deleterious psychological and physiological states, arousal has been at the center of several classic theories in social psychol-

ogy as the proposed mediator of behavioral outcomes. From social facilitation to cognitive dissonance, arousal has been implicated in both positive and negative performance outcomes and psychological states.

For several decades social psychologists have theorized that how one construes bodily responses, such as arousal, can affect behavior, emotions, and even performance (e.g., Niedenthal, 2007; Schachter & Singer, 1962). The notion that construal has important behavioral consequences downstream is also consistent with contemporary models of emotion like Gross's (1998) emotion regulation model and Barrett's core affect theory (2006). In the latter theory, Barrett and colleagues argue that the conceptualization process transforms internal states into meaningful psychological states by integrating bodily changes with external sensory information and situation specific knowledge. For example, high arousal might be interpreted as fear or excitement depending on a variety of factors including knowledge of the situation, context, and experience.

More specifically with regards to reappraisal, Gross argues that appraisal processes occur early in the emotion-generative process, and the downstream outcome (the experienced emotion) is most easily altered by changing appraisals of the meaning of internal states (Gross, 2002). In a recent paper examining the effects of reappraisal on cardiovascular responses, Mauss, Cook, Cheng, and Gross (2007) found that participants high in reappraisal tendencies had stronger sympathetic activation (i.e., interpreted as a *challenge* response) during a mental arithmetic task combined with an anger provocation than those low in reappraisal. They interpreted these responses as reappraisal resulting in an adaptive affective profile.

* Corresponding author. Address: Harvard University, 33 Kirkland Street, WJH 1420, Cambridge, MA 02138, United States.

E-mail addresses: wbm@wjh.harvard.edu (W.B. Mendes), tschmader@psych.ubc.ca (T. Schmader).

In an academic context, students, by default, seem to appraise arousal during a high-stakes test as an indication of anxiety that will be detrimental for performance (Johns, Inzlicht, & Schmader, 2008). Therefore, encouraging test-takers to reappraise arousal as a beneficial, promotive state may help break the association between arousal and anxiety, which should then improve performance.

Physiologically, arousal is associated with increases in sympathetic nervous system (SNS) responses, which can be measured by examining catecholamine levels. Increased SNS activity has been associated with two distinct motivational states: challenge and threat (Blascovich & Mendes, 2000), with challenge states typically resulting in relatively greater SNS activation. Unlike threat, challenge is also characterized by performance improvement, which is consistent with the strong linear relationship noted between catecholamine levels and cognitive performance (see Dienstbier, 1989). Challenge states have been routinely linked to better cognitive performance in a variety of domains including pattern-detection, cooperative games, and decision-making tasks (Blascovich, Mendes, Hunter, & Salomon, 1999; Kassam, Koslov, & Mendes, in press; Mendes, Major, McCoy, & Blascovich, 2008). Thus, test-takers would presumably be at an advantage if they appraised arousal as a challenge signal, rather than a threat signal during test performance.

In this research, appraisals of arousal were manipulated and GRE performance was measured both in and outside the laboratory. sAA levels, a non-specific measure of SNS activation, were measured prior to testing in the laboratory session. We expected reappraisal participants to exhibit increased sAA levels (indexing relatively more engagement and challenge orientation), as well as better GRE performance, compared to controls. Then, if the appraisal manipulation generalized to actual GRE testing situations, the GRE scores of participants told to reappraise their arousal should also exceed those of controls. Although some recent evidence suggests that dispositional differences in reappraisal tendencies predicts SNS activation and performance (Schmader, Forbes, Zhang, & Mendes, 2009), no prior research has experimentally manipulated appraisal processes to investigate the effects on actual test performance. Such evidence is critical to advance our understanding of causal mechanisms necessary to design successful intervention strategies.

Methods

Participants

Sixty students (31 male, 29 female) planning to take the GRE within 3 months were initially recruited and scheduled for a laboratory session. Of these 60 participants, 28 (57% male) actually took the GRE in the required time window and returned to the lab for the follow-up session.¹ Thus, all participants in the final sample were preparing to take the GRE, completed preparation material, and took the GRE test within 3 months of the laboratory session.

Procedures

Participants were initially scheduled for two lab visits on consecutive days. The first visit lasted less than 30 min and allowed us to obtain a saliva sample (T0) that indexed sAA levels on a con-

trol day. At the same time the following day, participants reported back to the lab for the practice GRE. This visit lasted 2.5 h. On both days participants were instructed to refrain from caffeine and strenuous exercise for at least 2 h prior to arrival. For the practice GRE, we created a similar testing environment that participants would experience during the actual GRE. Each participant was seated at a computer, given scratch paper, and instructed to “try as hard on today’s practice test as you will during the actual GRE.”

After consent and collection of the practice day saliva sample (T1), participants received GRE test instructions, which included the reappraisal manipulation. Participants in both conditions first heard/read the following instructions:

“The goal of this research is to examine how physiological arousal during a test correlates with performance. Because it is normal for people to feel anxious during standardized tests, the saliva samples . . . will be analyzed for hormones that indicate your arousal level.”

Although the cover story for the study ended here for control participants, those assigned to the reappraisal condition then heard/read:

“People think that feeling anxious while taking a standardized test will make them do poorly on the test. However, recent research suggests that arousal doesn’t hurt performance on these tests and can even help performance. . . people who feel anxious during a test might actually do better. This means that you shouldn’t feel concerned if you do feel anxious while taking today’s GRE test. If you find yourself feeling anxious, simply remind yourself that your arousal could be helping you do well.”

After instructions, participants were given practice problems and then completed quantitative and verbal sections from a GRE practice test (each scored 200–800), with order counterbalanced. Prior to testing, but after manipulations and practice problems, we obtained a second saliva sample (T2) to assess SNS activation.² After collection, saliva samples were stored in a –80° C freezer until they were sent on dry ice to Dresden, Germany. There they were thawed and centrifuged at 3000 rpm for 5 min. Concentration of sAA was measured by an enzyme kinetic method (α -amylase; Roche Diagnostics).

Participants reported back to the lab 1–3 months later. During this visit participants provided a copy of their Educational Testing Service (ETS) score report and completed GRE experience questionnaires, which assessed the amount of arousal experienced during the testing session, whether participants believed arousal helped or hurt performance, how much they worried about feeling anxious, and how confident participants were of themselves during testing, all on 7-point scales.

Results

We first examined if there were pre-existing differences between our manipulated conditions. No differences emerged in SAT scores or college GPA as a function of sex or appraisal, $ps > .40$. We also examined sAA levels on the control day and baseline (T0 and T1) and found no differences, $ps > .20$. Finally, order of the practice test sections did not influence the results.

Practice test

Practice GRE performance

GRE scores were analyzed in a 2 (appraisal: reappraisal vs. control) \times 2 (section: quantitative vs. verbal) ANOVA with appraisal as

¹ Our 50% attrition rate was due to participants who did not take the GRE during the time allotted. We compared those who did take the GRE to those who did not and there were many differences. For example, those who did not take the GRE compared to those who did had lower GPAs, and practiced less for the GRE before the laboratory session, $ps < .05$. However, importantly for this study, those who did take the GRE did not differ on any of these dimensions by reappraisal condition.

² sAA peaks within minutes of an event. Thus, we timed T1 collection so that it was constant across participants (Nater et al., 2005).

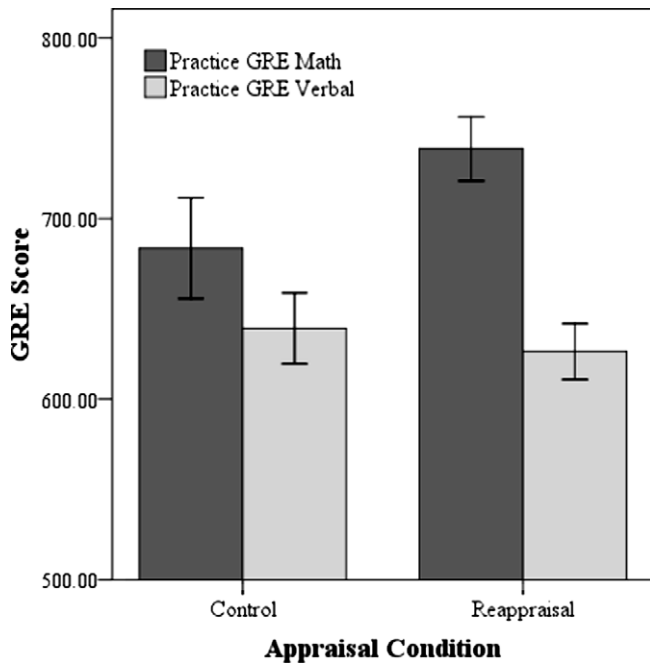


Fig. 1. Practice GRE performance as a function of appraisal condition and test section. Scores could range from 200 to 800. Error bars represent \pm standard error of the mean.

a between subjects factor and section within subjects. We observed a marginal Appraisal \times Section interaction, $F(1, 26) = 3.30$, $p = .081$, $d = .71$. Contrasts (Kirk, 1995) indicate that reappraisal participants performed significantly better ($M = 738.57$, $SD = 66.43$) than controls ($M = 683.57$, $SD = 104.63$) on the math section, $F(1, 26) = 4.35$, $p = .047$, $d = .82$, whereas no differences emerged on the verbal section, $F < 1$ (see Fig. 1). These data demonstrate that participants told to reappraise arousal experienced performance facilitation.

Sympathetic reactivity

sAA reactivity was computed as a difference score: sAA levels taken at T1 were subtracted from those measured at T2. As shown in Fig. 2, reappraisal participants exhibited a significant increase in sAA levels ($M = 24.86$, $SD = 37.58$) compared to controls ($M = -10.66$, $SD = 35.66$), $t(26) = 2.57$, $p = .016$, $d = 1.01$. Furthermore, a related-samples t -test indicates that reappraisal participants' sAA levels increased significantly from their baseline, $t(13) = 2.49$, $p = .027$, $d = .97$ (95% CI = 10.94–38.78), whereas controls showed a non-significant change, $p > .25$. Thus, reappraisal led to a large increase in SNS activity immediately preceding testing, whereas the control condition showed no changes. This is consistent with the idea that challenge or approach states are characterized by greater SNS activation.³ We re-ran all analyses controlling for individual differences in sAA using T0 levels, which were obtained on the control day and less likely to be influenced by pre-exam anxiety. All effects persisted.

We then examined the relationship between sAA reactivity and performance for math and verbal sections separately. For control participants, there was no association between sAA levels and performance on either section, $ps > .30$. However, among those assigned to the reappraisal condition, increases in sAA were related

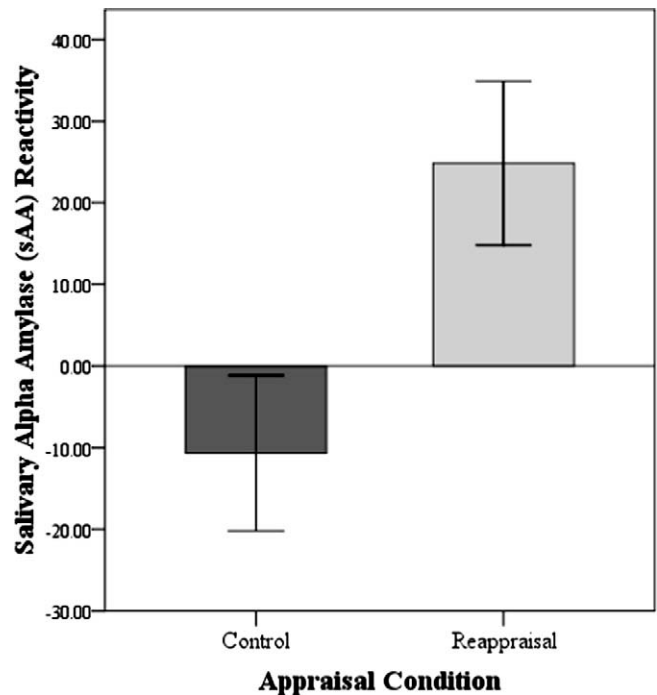


Fig. 2. Salivary alpha amylase (sAA) reactivity in response to the appraisal manipulation. Positive values indicate an increase in sAA levels from baseline to post appraisal manipulation. Error bars represent \pm standard error of the mean.

to better math performance, $r = .57$, $p = .033$. sAA levels were not related to verbal scores in the reappraisal condition, $p > .80$.

Post GRE lab visit

Subjective experience

We then examined participants' experiences after taking the actual GRE (Table 1). These data suggested that the reappraisal manipulation persisted over time. Compared to controls, reappraisal participants reported that arousal helped performance more, $t(26) = 2.53$, $p = .018$, $d = .99$; worried less about feeling anxious, $t(26) = 1.70$, $p = .102$, $d = .67$; and reported feeling less unsure of themselves, $t(25) = 2.46$, $p = .022$, $d = .97$. Thus, the laboratory manipulation generalized to the actual GRE testing session.

Actual GRE performance

GRE scores from participants' ETS reports⁴ were then analyzed. An Appraisal \times Section interaction replicated the effect observed in the laboratory, $F(1, 26) = 5.20$, $p = .031$, $d = .89$. As shown in Fig. 3, the appraisal manipulation had no effect on verbal performance, $F < 1$. However, reappraisal participants performed significantly better ($M = 770.00$, $SD = 63.64$) than controls on the math section ($M = 705.71$, $SD = 93.37$), $F(1, 26) = 6.85$, $p = .015$, $d = 1.03$. Thus, the appraisal manipulation facilitated performance during actual GRE testing. Moreover, sAA reactivity measured during the laboratory session predicted appraisal participants' actual GRE-math performance ($\beta = .492$, $p = .008$) suggesting that participants with the greatest activation to the reappraisal manipulation in the lab may have also had the largest SNS activation during the actual testing situation, thus reaping the greatest benefits from the reappraisal manipulation.

³ The absence of a significant sAA reactivity effect in the control condition suggests that these participants were either less engaged or threatened relative to reappraisal participants. sAA reactivity, alone, is not sufficient to differentiate between these two alternatives.

⁴ We also analyzed percentile scores because the percentile rank corresponding to raw GRE score varies from test to test. Analysis of percentiles did not impact the results in any way.

Table 1

Responses to GRE test experience questionnaire items taken after participants completed the actual GRE test (7-point scale).

Questionnaire item	Appraisal condition			
	Reappraisal		No appraisal	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Does arousal help performance?	3.86 _a	1.51	2.57 _b	1.16
Were you worried about feeling anxious?	2.46 _a	1.43	3.54 _a	1.86
Were you unsure of your performance?	2.75 _a	1.36	4.39 _b	1.89

Note: Different subscript letters indicate significant mean differences within the row.

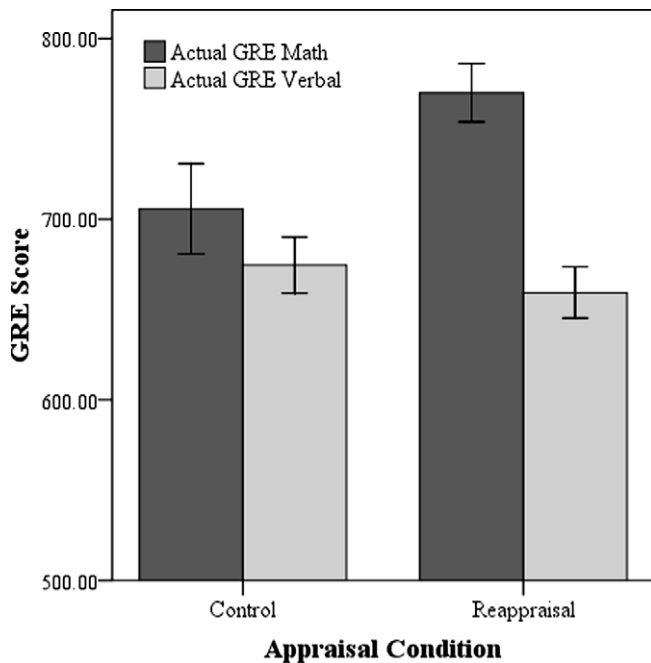


Fig. 3. Actual GRE performance as a function of appraisal condition and test section. Scores could range from 200 to 800. Error bars represent \pm standard error of the mean.

We re-ran all performance analyses using available covariates of academic performance (SAT, GPA, prior coursework, time spent studying for the GRE), and the effects of reappraisal persisted with these covariates. Furthermore, the length of time between the laboratory session and actual GRE testing did not moderate these effects.

Discussion

This study examined the effect of reappraising arousal as a *challenge* response on GRE performance both in the laboratory and in actual testing situations. During the laboratory session, participants who were instructed that arousal signaled good performance exhibited elevated catecholamine levels and performed better on the GRE-math section compared to controls. The data from the actual GRE test replicated the pattern of performance produced in the laboratory, suggesting that manipulations of reappraisal can generalize outside of the laboratory to real-world testing situations and influence test-takers' scores.

It may seem remarkable that a reappraisal manipulation given over a month before participants took the GRE was sufficient to improve performance. But similarly, a simple writing exercise inter-

vention given at the start of an academic term improved final grades by 40% (Cohen, Garcia, Apfel, & Master, 2006), and persisted to benefit grade point averages two years after the initial manipulation (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). In the current research, our laboratory manipulation of arousal appraisal appeared to have lasting effects, both subjectively and behaviorally, one to three months after the manipulation.

In this study reappraising arousal improved math performance, but had no effect on verbal performance, nor was arousal associated with verbal scores in the control condition ($p > .40$). This lack of consistency may result from characteristics of the problems found in each section. Math problems generally require test-takers to use executive resources to actively process and compute information, whereas the verbal section is dominated by problems (e.g., antonyms and analogies) requiring the retrieval of information from long-term storage with fewer active processing requirements (e.g., Halpern, 2004). Since research suggests that reappraising arousal improves executive functioning (Johns et al., 2008), it may not be surprising that appraisal improved only math performance. Additionally, our reappraisal manipulation could have facilitated performance on the math, but not the verbal, section because students tend to exhibit more negative implicit attitudes toward math domains relative to verbal domains (Nosek, Banaji, & Greenwald, 2002).

Importantly, our findings are consistent with previous work on the effects of SNS activation, specifically catecholamine reactivity, on cognitive performance. Across several large scale studies, greater catecholamine increases were associated with better task performance (see Dienstbier (1989) for a review). Indeed, in Dienstbier's seminal review he concluded that "despite the high difficulty level of ... these tasks ... these data indicate *no curvilinear relations*; naturally evoked peripheral catecholamines *never seem to be too high* for optimal performance" (italics added, p. 86).

Although this research demonstrates that reappraisal improves math performance beyond the laboratory, the naturalistic setting makes it difficult to isolate mechanisms. On the one hand, appraisal might have improved performance by increasing study time if arousal cued an approach orientation towards the test, rather than avoidance. Alternatively, reappraisal participants may have remembered and reinterpreted the arousal they were feeling on the day of the actual GRE test. Future research may seek to specify mechanisms.

In sum, these findings show that people's appraisals of their internal states are flexible. As such, the manner in which internal states are interpreted can have profound effects on emotions, physiology, and behavior. In this research, we focused on the effects of reappraisal of arousal on GRE performance. However, the data presented here can be applied beyond standardized testing. For instance, if students construe criticism from professors as constructive (challenge) rather than derisive (threat), it could help improve performance. Thus, this research suggests that our physiology and behavior may be strongly dependent on our cognitive appraisals of internal states.

Acknowledgments

We thank Modupe Akinola and Kristen Lindquist for their helpful comments. This research was funded in part by an NHLBI Grant (R01 HL079383) (WBM) and an NIMH Grant (R01 MH071749) (TS).

References

- Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, 10, 20–46.
- Blascovich, J. (1992). A biopsychosocial approach to arousal regulation. *Journal of Social and Clinical Psychology*, 11, 213–237.

- Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition*. Cambridge, UK: Cambridge University Press.
- Blascovich, J., Mendes, W. B., Hunter, S. B., & Salomon, L. (1999). Social "facilitation" as challenge and threat. *Journal of Personality and Social Psychology, 77*, 68–77.
- Cassaday, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27*, 270–295.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social–psychological intervention. *Science, 313*, 1307–1310.
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*, 400–403.
- Dienstbier, R. A. (1989). Arousal and physiological toughness: Implications for mental and physical health. *Psychological Review, 96*, 84–100.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271–299.
- Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*, 281–291.
- Halpern, D. F. (2004). A cognitive-process taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science, 13*, 135–139.
- 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*, 691–705.
- Kassam, K. S., Koslov, K., & Mendes, W. B. (in press). Decisions under distress: Stress profiles influence anchoring and adjustment. *Psychological Science*.
- Kirk, R. (1995). *Experimental design*. Pacific Grove, CA: Brooks/Cole Publishing Company.
- Mauss, I., Cook, C. L., Cheng, Y. J., & Gross, J. J. (2007). Individual differences in cognitive reappraisal experiential and physiological responses to anger provocation. *International Journal of Psychophysiology, 66*, 116–124.
- Mendes, W. B., Major, B., McCoy, S., & Blascovich, J. (2008). How attributional ambiguity shapes physiological and emotional responses to social rejection and acceptance. *Journal of Personality and Social Psychology, 94*, 278–291.
- Nater, U. M., Rohleder, N., Gaab, J., Berger, S., Jud, A., Kirschbaum, C., et al. (2005). Human salivary alpha-amylase reactivity in a psychosocial stress paradigm. *International Journal of Psychophysiology, 55*, 333–342.
- Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*, 1002–1005.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology, 83*, 44–59.
- Schachter, S., & Singer, J. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review, 69*, 379–399.
- 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.