Title

Cytokine responses and math performance: The role of stereotype threat and anxiety reappraisals

Authors

John-Henderson, Neha A
Rheinschmidt, Michelle L
Mendoza-Denton, Rodolfo

Publication Date

2015

DOI

10.1016/j.jesp.2014.10.002

Peer reviewed

Journal of Experimental Social Psychology 56 (2015) 203–206

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

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

FlashReport

Cytokine responses and math performance: The role of stereotype threat and anxiety reappraisals

Neha A. John-Henderson *, Michelle L. Rheinschmidt, Rodolfo Mendoza-Denton

Department of Psychology, University of California, Berkeley, USA

HIGHLIGHTS

- Female college students take a math exam described as gender-fair or gender-biased.
- In one condition, participants directed to reappraise physiological arousal.
- Performance on math exam and post-exam levels of the cytokine IL-6 were measured.
- Reappraisal of physiological arousal buffers inflammatory responses to exam across conditions.
- Reappraisal of arousal especially effective buffer of inflammatory responses in stereotype threat condition.

ARTICLE INFO

Article history:
Received 11 January 2014
Revised 1 October 2014
Available online 30 October 2014

Keywords:
Stereotype threat
Reappraisal
Stress
Gender
Inflammation

ABSTRACT

This research independently manipulated two potential attenuators of stereotype threat – reappraisal of anxiety and test framing – to explore their independent and combined effects. Female participants took a difficult math exam that was described as gender-biased or gender-fair and were told that anxious arousal could positively impact performance or were given no information regarding arousal. Levels of the cytokine Interleukin-6 (IL-6), an immune marker of inflammation, were measured in oral mucosal transudate (OMT) both before and after the exam. Our findings indicate that directing reappraisal of physiological arousal attenuated increases in IL-6 across test framing conditions, and was especially effective under stereotype threat (i.e., gender-biased test condition). Reappraisal also mapped onto better test performance in the threat condition. Together, these findings provide insight into the unique and interactive effects of two situational interventions meant to reduce stereotype threat, indexed here by both physiological and performance-based correlates of threat.

Introduction

A widely shared stereotype in our society alleges that women possess weaker mathematical abilities than men (Cheryan, Plaut, Davies, & Steele, 2009; Swim, 1994). Women risk being judged by this stereotype and worry that they will confirm negative stereotypes when performing math tasks; such gender-based stereotype threat has been linked to anxiety, cognitive resource depletion, and underperformance (e.g., Johns, Schmader, & Martens, 2005; Mendoza-Denton, Kahn, & Chan, 2008).

Prior research has shown that reappraising anxious arousal as beneficial to performance yields better performance and more adaptive physiological responses relative to providing no reappraisal instructions (Jamieson, Mendes, Blackstock, & Schmader, 2010; Jamieson, Nock, & Mendes, 2012). Targets of negative stereotypes may be especially likely to benefit from reappraisal, given the heightened state of anxiety and physiological arousal inherent to the experience of stereotype threat (Schmader, Johns, & Forbes, 2008). Indeed, Johns, Inzlicht, and Schmader (2008) observed benefits of reappraising anxiety on test performance among participants under stereotype threat; however, to our knowledge no research has addressed whether reappraisal strategies also reduce the physiological consequences associated with stereotype threat. We fill this gap in the literature by experimentally testing whether reappraisal of anxiety, specifically among women in math contexts, diminishes both performance impairments and increases in a marker of inflammation associated with stereotype threat.

In this research we specifically focused on the effects of reappraisal on the pro-inflammatory cytokine Interleukin-6 (IL-6). Pro-inflammatory cytokines such as IL-6 are critical for orchestrating the body's inflammatory response, which is crucial to fighting injury or infection (Parkin & Cohen, 2001; Segerstrom & Miller, 2004). However, if the inflammatory response becomes persistent or exaggerated it can lead to a host of diseases and health conditions. Inflammation is increasingly recognized as a risk factor for illnesses such as cardiovascular disease, autoimmune

* Corresponding author at: Department of Psychology, University of California, Berkeley, 3210 Tolman Hall #1650, Berkeley, CA 94720-1650, USA.
E-mail address: nehajohn@berkeley.edu (N.A. John-Henderson).

http://dx.doi.org/10.1016/j.jesp.2014.10.002
0022-1031/

disorders, and some cancers (Cesari, Penninx, & Newman, 2003; Nathan & Ding, 2010).

Prior research (e.g., Dickerson, Gable, Irwin, Aziz, & Kemeny, 2009; Murali, Hanson, & Chen, 2007) suggests shifts in inflammatory cytokine levels in response to situational stressors (e.g., taking an exam). While IL-6 in particular can exert both inflammatory and anti-inflammatory effects (Scheller, Chalaris, Schmidt-Arras, & Rose-John, 2011), prior research characterizes increases in IL-6 specifically in response to a stressor as signaling an inflammatory response (Dickerson et al., 2009; John-Henderson, Rheinschmidt, Mendoza-Denton, & Francis, 2014; Slavich, Way, Eisenberger, & Taylor, 2010). In this research, we assessed changes in IL-6 in oral mucosal transudate (OMT). While levels of inflammatory cytokines in OMT are not a reflection of systemic inflammation, which is assessed through blood (Fernandez-Botran, Miller, Burns, & Newton, 2011), inflammatory markers in OMT have been shown to be affected by acute situational stressors (Chiang, Eisenberger, Seeman, & Taylor, 2012; John-Henderson et al., 2014; Slavich et al., 2010).

Method

Participants and procedure

Ninety-seven female undergraduates at UC Berkeley participated for partial course credit. Their ages ranged from 18 to 35 years ($M = 20.88$, $SD = 2.77$), and the ethnic composition was 53.6% Asian, 21.6% White, 16.4% Hispanic, .06% other, and .02% African American.

This study crossed previously published manipulations of stereotype threat (Spencer, Steele, & Quinn, 1999) and reappraisal of anxiety (Jamieson et al., 2010). We manipulated stereotype threat by following the exact procedure of Spencer et al. (1999). More specifically, we described the math exam as a test of intellectual ability for solving math problems that had or had not produced gender differences in performance. We also varied whether participants received reappraisal instructions or no information about reappraisal. Using the wording of Jamieson et al. (2010), the reappraisal instructions encouraged participants to view arousal and anxiety as helpful to test performance. Participants were randomly and independently assigned to condition. Participants provided samples of oral mucosal transudate (OMT) at three time points (baseline, post-exam, and recovery) to assess changes in IL-6 in response to the experimental manipulations. The experimenters were blind to participants' condition, as all manipulations occurred via paper-based instructions embedded in the math examination packet. Participants were given 30 min to complete the test.

Measures

Test performance

We measured the number of correct responses on a 17-item math exam modeled after the Graduate Record Examination ($M = 11.6$, $SD = 4.3$).

Inflammation measures

We assessed IL-6 levels in OMT. Participants provided a baseline OMT sample for IL-6 measurement upon arrival to the lab ($M = .45$ pg/mL, $SD = .30$). An OraSure collection device (Epitope, Beaverton, OR) was placed between the lower cheek and gum for 2 min. The exam lasted for 30 min, after which a second sample of OMT was taken using the same method to assess changes in IL-6 ($M = 1.98$ pg/mL, $SD = 2.52$). A third, final sample was taken 30 min after the second sample to gauge recovery ($M = 1.82$ pg/mL, $SD = 2.44$).

All OMT samples were immediately frozen and stored at −80°°C. IL-6 concentrations were determined by an enzyme-linked immunosorbent assay using commercially available kits (R&D systems, Minneapolis, MN). The intra-assay coefficient of variation (CV) was 6.5% and the inter-assay CV was 8.7%. Shapiro–Wilk tests (Shapiro & Wilk, 1965) revealed that IL-6 values at each time point were not normally distributed (baseline: $W = .91$, $p < .001$, post-exam: $W = .67$, $p < .001$, recovery: $W = .65$, $p < .001$). Thus, following prior research (John-Henderson, Jacobs, Mendoza-Denton, & Francis, 2013) we added a constant of one to all raw values (see Osborne, 2002) before applying a log-transformation to IL-6 values. In addition to measuring IL-6 levels, we measured levels of total protein in each OMT sample using the BCA protein assay with bovine serum albumin as the standard (Thermoscientific, Rockford, IL) to control for individual differences in salivary flow rate (Dickerson, Kemeney, Aziz, Kim, & Fahey, 2004). Salivary flow rate was specific to each sample timepoint, given documented fluctuations in these rates under acute stress (Bakke et al., 2004). All total protein samples were run in triplicate following kit instructions.

Body mass index (BMI)

Given its relationship with levels of IL-6 in previous research (Khaodhiar, Ling, Blackburn, & Bistrian, 2004), we calculated participants' BMI ($M = 21.96$, $SD = 3.10$) from their self-reported height and weight for use as a covariate in the IL-6 analyses.

Results

IL-6 reactivity

We conducted a general linear model analysis predicting IL-6 levels as a function of the sample timepoint (3 within-participant levels: baseline, post-exam, recovery) and our between-participant factors of test framing (2 levels: gender-biased, gender-fair) and reappraisal instructions (2 levels: instructions, no instructions). We contrast-coded condition assignments prior to analyses (for reappraisal instructions: −.5 = no mention of arousal, .5 = arousal helps performance; for test framing: −.5 = gender-fair, .5 = gender-biased). We included total protein levels at each timepoint and BMI as covariates; all were mean-centered. Based on Mauchly's test of sphericity, $\chi^2(2) = 42.79$, $p < .001$, we applied a Greenhouse–Geisser correction ($\varepsilon = .72$) to estimate our overall effect more conservatively. We report adjusted degrees of freedom below.

Our analysis revealed the predicted 3-way interaction of time, test framing, and reappraisal instructions, $F(1.43, 121.50) = 10.25$, $p < .001$ $\eta_p^2 = .11$.[1] BMI accounted for a significant amount of variance between participants, $F(1, 85) = 4.39$, $p = .04$, $\eta_p^2 = .05$, but total protein levels at baseline, post-exam, and recovery did not, p's > .34, all $\eta_p^2 < 01$. We broke down the 3-way interaction by looking at the effects of reappraisal instructions within test framing condition. Specifically, we examined whether changes in IL-6 could be characterized by linear and/or quadratic trends over time as a function of reappraisal condition.

In the gender-biased test condition, we observed a significant simple interaction between time and reappraisal instruction condition $F(1, 85) = 59.75$, $p < .001$, $\eta_p^2 = .41$, for the linear effect of time, and $F(1, 85) = 65.32$, $p < .001$, $\eta_p^2 = .44$, for the quadratic effect of time. We conducted simple comparisons at each timepoint to identify significant differences in IL-6 levels by reappraisal condition, applying a Bonferroni correction to reduce Type 1 error. As illustrated in Panel A of Fig. 1, these comparisons indicate no IL-6 differences at baseline, $F(1, 85) = 1.17$, $p = .28$, $\eta_p^2 = .01$. However, women in the reappraisal instruction condition (vs. no instructions condition) had significantly lower levels of IL-6 during the post-exam, $F(1, 85) = 61.30$, $p < .001$, $\eta_p^2 = .42$, and recovery periods, $F(1, 85) = 54.15$, $p < .001$, $\eta_p^2 = .39$.

In the gender-fair test condition, we observed a significant simple interaction between time and the reappraisal instruction manipulation, $F(1, 85) = 6.85$, $p = .01$, $\eta_p^2 = .08$, for the linear effect of time, and $F(1, 85) = 6.32$, $p = .01$, $\eta_p^2 = .07$, for the quadratic effect of time.

[1] We confirmed that our models hold when using age, subjective general health, current illness (1 = yes; 0 = no), and chronic health conditions (1 = yes; 0 = no) as covariates. Further, we repeated the analyses while excluding anyone with a physical illness and/or chronic health condition, and our results hold.

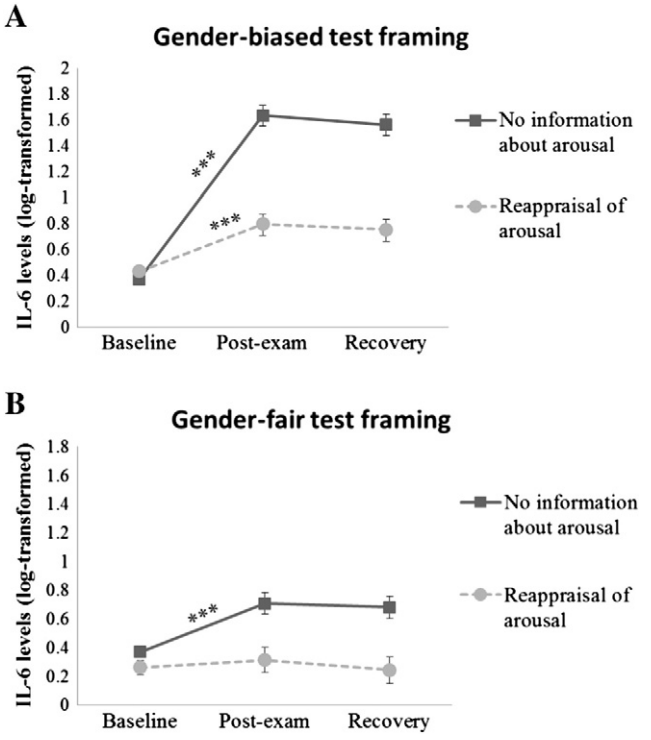

Fig. 1. Average IL-6 levels by time of sample (x-axis), reappraisal instructions (separate lines), and test framing (Panels A vs. B). IL-6 values have been transformed by adding a constant of 1 and then performing a log-transformation. Slopes between timepoints are not statistically significant ($p > .05$) unless noted above the slope. All p values were adjusted with a Bonferroni correction. Error bars represent standard error. *** $p < .001$.

We isolated these divergent trajectories by conducting simple comparisons at each timepoint, again using a Bonferroni correction (see Fig. 1, Panel B). We found no statistically significant IL-6 differences at baseline according to reappraisal instruction condition, $F(1, 85) = 2.91$, $p = .09$, $\eta_p^2 = .03$. However, at the post-exam and recovery timepoints, receiving reappraisal instructions yielded lower levels of IL-6 relative to receiving no instructions, $F(1, 85) = 10.83$, $p = .001$, $\eta_p^2 = .11$, for post-exam, and $F(1, 85) = 12.66$, $p = .001$, $\eta_p^2 = .13$, for recovery.

Looking within women who did not receive reappraisal instructions, we observed a significant simple interaction between time and test framing condition, $F(1, 85) = 57.58$, $p < .001$, $\eta_p^2 = .40$, for the linear effect of time, and $F(1, 85) = 64.16$, $p < .001$, $\eta_p^2 = .43$, for the quadratic effect of time. Women in the gender-biased condition had higher levels of IL-6 at post-exam, $F(1, 85) = 71.23$, $p < .001$, $\eta_p^2 = .46$, and recovery, $F(1, 85) = 60.93$, $p < .001$, $\eta_p^2 = .42$, than those in the gender-fair condition.

Among women who received reappraisal instructions, we observed a significant simple interaction between time and testing framing condition, $F(1, 85) = 6.90$, $p = .01$, $\eta_p^2 = .08$, for the linear effect of time, and $F(1, 85) = 6.74$, $p = .01$, $\eta_p^2 = .07$, for the quadratic effect of time. However, we found smaller effects of test framing at post-exam, $F(1, 85) = 15.65$, $p < .001$, $\eta_p^2 = .16$, and recovery, $F(1, 85) = 16.47$, $p < .001$, $\eta_p^2 = .16$, than we did in the no reappraisal condition.

In summary, across both test framing conditions, women who received reappraisal instructions (vs. no instructions) showed lower levels of IL-6 following a difficult math exam. We observed larger effects of reappraisal in the gender-biased condition. Further, we observed larger effects of stereotype threat when no reappraisal instructions were given. In the gender-biased condition, women who did not receive reappraisal instructions had higher peak levels of IL-6 than all other women (see Fig. 1), suggesting that these women were especially vulnerable to rises in IL-6 following the exam.

Test performance

We observed a significant zero-order correlation between post-exam inflammation levels and exam performance, $r(97) = -.49$, $p < .001$. We observed a non-significant partial correlation between post-exam inflammation levels and exam performance, after controlling for our experimental conditions, BMI, IL-6 levels at baseline, and total protein levels at baseline and post-exam, $r(86) = -.08$, $p = .46$.

We conducted an ANOVA with test framing condition and reappraisal instruction condition as independent variables to see whether our experimental manipulations interacted to predict test performance. We observed main effects of both the test framing manipulation, $F(1, 93) = 47.38$, $p < .001$, $\eta_p^2 = .34$, and the reappraisal instruction manipulation, $F(1, 93) = 9.34$, $p = .003$, $\eta_p^2 = .09$. These main effects were qualified by a marginally significant interaction between the two manipulations, $F(1, 93) = 3.34$, $p = .07$, $\eta_p^2 = .04$. As illustrated in Fig. 2, participants performed significantly better under stereotype threat when given reappraisal instructions versus no instructions, $F(1, 93) = 12.05$, $p = .001$, $\eta_p^2 = .12$. Women's performance in the gender-fair test framing condition did not differ as a function of reappraisal instructions, $F(1, 93) = .75$, $p = .39$, $\eta_p^2 = .008$. Thus, reappraisal was most beneficial to performance under stereotype threat.

We also tested whether our stereotype threat manipulation was more potent in the absence of reappraisal instructions. Looking at women who did not receive reappraisal instructions, we observed worse performance in the gender-biased condition compared to the gender-fair condition, $F(1, 93) = 38.33$, $p < .001$, $\eta_p^2 = .29$. This difference by test framing condition was indeed less pronounced among women who received reappraisal instructions, $F(1, 93) = 12.65$, $p = .001$, $\eta_p^2 = .12$.

General discussion

Our experimental findings provide evidence that reappraisal of anxiety yields better performance and lower IL-6 activation under stereotype threat relative to receiving no reappraisal instructions. We extend Jamieson et al. (2010) by exploring the benefits of reappraisal among negatively stereotyped individuals, who have known vulnerabilities to under-performance and increases in IL-6 under stereotype threat (John-Henderson et al., 2014).

We might have expected that reappraisal helps performance in any testing situation, given that reappraisal of anxiety has been found to help even members of non-stigmatized groups (e.g., men; Jamieson et al., 2010). Here, we did not find that reappraisal boosted performance relative to no reappraisal instructions when the test was framed as gender-fair. The discrepancy in our findings might be explained by the fact that Jamieson et al.'s (2010) research did not make stereotypes or gender differences salient. By contrast, our findings suggest that a

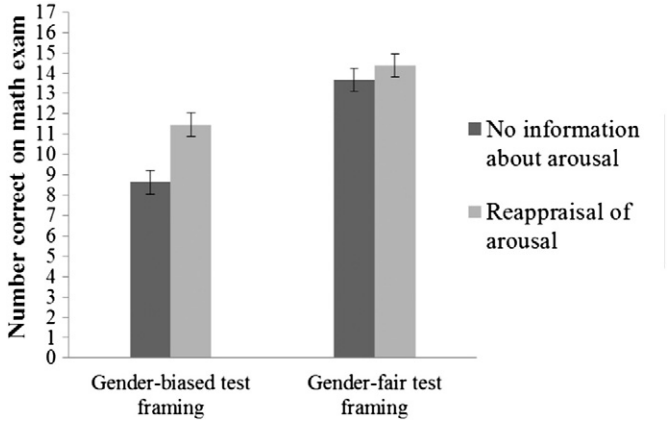

Fig. 2. Performance on a mathematics examination as a function of the reappraisal instruction condition and the test framing condition. Error bars represent standard error.

gender-fair message may be enough, even without reappraisal, to release women from the cognitive intrusions associated with stereotype threat.

Limitations and future directions

A limitation of our research is that we did not inquire about the smoking behavior of our participants, a possible predictor of oral inflammation[2]. Given our findings, it will be important for future research to examine whether the patterns observed here emerge via measurement of inflammation in blood plasma, so as to examine more directly the implications of this research for longer-term health outcomes and vulnerabilities.

Our findings might also align with work showing increases in salivary alpha amylase (sAA), reflecting increased sympathetic nervous system activation when reappraisal instructions are given in testing contexts (Jamieson et al., 2010). In addition, based on a large body of literature showing that cortisol levels increase in response to naturalistic acute stressors (e.g., Weekes et al., 2006), we would expect to find that changes in cortisol would parallel the increase in IL-6 in this research. An important extension of this work will be to examine the simultaneous or differential trajectories of these physiological systems. This type of multi-modal investigation would allow for a more thorough understanding of the mechanisms that explain the variance in IL-6 levels documented here.

Acknowledgments

The first two authors contributed equally to the development of this project and the ideas within it. We are grateful for the data collection assistance of Leeran Baraness, Hardev Chhokar, Mary Cunningham, and Meital Mashash. We would also like to thank Darlene Francis for her feedback during initial stages of the project and for her support in conducting the assays for this research. This work was supported by the National Science Foundation Graduate Research Fellowship awarded to Michelle L. Rheinschmidt and by the National Science Foundation Award HRD-1306709.

References

American College Health Association (2013). American College Health Association-National College Health Assessment (ACHA-NCHA). Available from. http://www.acha-ncha.org/

Bakke, M., Tuxen, A., Thomsen, C.E., Bardow, A., Alkjaer, T., & Jensen, B.R. (2004). Salivary cortisol level, salivary flow rate and masticatory muscle activity in response to acute mental stress: A comparison between aged and young women. *Gerontology*, *50*, 383–392. http://dx.doi.org/10.1159/000080176.

Cesari, M., Penninx, B.W., & Newman, A.B. (2003). Inflammatory markers and cardiovascular disease (The Health, Aging and Body Composition [Health ABC] Study). *American Journal of Cardiology*, *92*, 522–528. http://dx.doi.org/10.1016/S0002-9149(03)00718-5.

Cheryan, S., Plaut, V.C., Davies, P.G., & Steele, C.M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*(6), 1045–1060. http://dx.doi.org/10.1037/a0016239.

Chiang, J.J., Eisenberger, N.I., Seeman, T.E., & Taylor, S.E. (2012). Negative and competitive social interactions are related to heightened proinflammatory cytokine activity. *Proceedings of the National Academy of Sciences*, *109*, 1878–1882. http://dx.doi.org/10.1073/pnas.1120972109.

Dickerson, S.S., Gable, S.L., Irwin, M.R., Aziz, N., & Kemeny, M.E. (2009). Social-evaluative threat and proinflammatory cytokine regulation: An experimental laboratory investigation. *Psychological Science*, *20*, 1237–1244. http://dx.doi.org/10.1111/j.1467-9280.2009.02437.x.

Dickerson, S.S., Kemeney, M.E., Aziz, N., Kim, K.H., & Fahey, J.L. (2004). Immunological effects of induced shame and guilt. *Psychosomatic Medicine*, *66*, 124–131.

Fernandez-Botran, R., Miller, J.J., Burns, V.E., & Newton, T.L. (2011). Correlations among inflammatory markers in plasma, saliva, and oral mucosal transudate in postmenopausal women with past intimate partner violence. *Brain, Behavior, and Immunity*, *25*, 314–321. http://dx.doi.org/10.1016/j.bbi.2010.09.023.

Jamieson, J.P., Mendes, W.B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology*, *46*, 208–212. http://dx.doi.org/10.1016/j.jesp.2009.08.015.

Jamieson, J.P., Nock, M.K., & Mendes, W.B. (2012). Mind over matter: Reappraising arousal improves cardiovascular and cognitive response to stress. *Journal of Experimental Psychology: General*, *141*, 417–422. http://dx.doi.org/10.1037/a0025719.

John-Henderson, N., Jacobs, E.G., Mendoza-Denton, R., & Francis, D.D. (2013). Wealth, health and the moderating role of implicit social class bias. *Annals of Behavioral Medicine*, *45*, 173–179. http://dx.doi.org/10.1007/s12160-012-9443-9.

John-Henderson, N., Rheinschmidt, M.L., Mendoza-Denton, R., & Francis, D.D. (2014). Performance and inflammation outcomes are predicted by different facets of SES. *Social Psychological and Personality Science*, *5*, 301–309. http://dx.doi.org/10.1177/1948550613494226.

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. http://dx.doi.org/10.1037/a0013834.

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. http://dx.doi.org/10.1111/j.0956-7976.2005.00799.x.

Khaodhiar, L., Ling, P.R., Blackburn, G.L., & Bistrian, B.R. (2004). Serum levels of interleukin-6 and C-reactive protein correlate with body mass index across the broad range of obesity. *Journal of Parental and Enteral Nutrition*, *28*, 410–415. http://dx.doi.org/10.1177/0148607 104028000410.

Mendoza-Denton, R., Kahn, K., & Chan, W. (2008). Can fixed views of ability boost performance in the context of favorable stereotypes? *Journal of Experimental Social Psychology*, *44*, 1187–1193. http://dx.doi.org/10.1016/j.jesp.2008.03.005.

Murali, R., Hanson, M.D., & Chen, E. (2007). Psychological stress and its relationship to cytokines and inflammatory disease. In N.P. Plotnikoff, R.E. Faith, & A.J. Murgo (Eds.), *Cytokines: Stress and Immunity* (pp. 29–49) (2nd ed.). Boca Raton, FL: Taylor & Francis Group.

Nathan, C., & Ding, A. (2010). Nonresolving inflammation. *Cell*, *140*, 871–882. http://dx.doi.org/10.1016/j.cell.2010.02.029.

Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation*, *8*(6) (Retrieved June 4, 2014 from http://PAREonline.net/getvn.asp?v=8&n=6).

Parkin, J., & Cohen, B. (2001). An overview of the immune system. *Lancet*, *357*, 1777–1789.

Scheller, J., Chalaris, A., Schmidt-Arras, D., & Rose-John, S. (2011). The pro- and anti-inflammatory effects of the cytokine Interleukin-6. *Biochimica et Biophysica Acta*, *1813*, 878–888. http://dx.doi.org/10.1016/j.bbamcr.2011.01.034.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336. http://dx.doi.org/10.1037/0033-295X.115.2.336.

Segerstrom, S.C., & Miller, G.E. (2004). Psychological stress and the immune system: A meta-analytic study of 30 years of inquiry. *Psychological Bulletin*, *130*, 601–630. http://dx.doi.org/10.1037/0033-2909.130.4.601.

Shapiro, S.S., & Wilk, M.B. (1965). An analysis of variance test for normality. *Biometrika*, *52*, 591–611.

Slavich, G.M., Way, B.M., Eisenberger, N.I., & Taylor, S.E. (2010). Neural sensitivity to social rejection is associated with inflammatory responses to social stress. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 14817–14822. http://dx.doi.org/10.1073/pnas.1009164107.

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.

Swim, J.K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, *66*, 21–36.

Weekes, N., Lewis, R., Patel, F., Garrison-Jakel, J., Berger, D.E., & Lupien, S.J. (2006). Examination stress as an ecological inducer of cortisol and psychological responses to stress in undergraduate students. *Stress*, *9*, 199–206. http://dx.doi.org/10.1080/1025389060102.

[2] Only 2.3% of female undergraduates on campus report using cigarettes 10 or more days per month (American College Health Association, 2013).