

## Supplemental Materials

### Pilot Study – Additional Results

#### SNS Reactivity during the Manipulation

In this study, we hypothesized that men and women would experience greater arousal during the ST task than the control task, as socially evaluative tasks elicit arousal for both men and women (Dickerson & Kemeny, 2004). However, to the extent that our ST manipulation evokes gender-specific threat, we hypothesized that only women would report greater demands relative to controls—characterizing a psychological state of threat.

Consistent with our hypotheses, participants in the ST condition showed significantly greater PEP reactivity during the task ( $M = -14.96$  milliseconds,  $SD = 14.23$  ms) than those in the control condition ( $M = -3.91$  ms,  $SD = 8.59$  ms;  $F(1, 141) = 30.85$ ,  $p < .001$ ;  $\eta_p^2 = .18$ ), adjusting for BMI given its influence on SNS activity (Carroll, Phillips, & Der, 2008). There was no main effect of gender on PEP reactivity ( $F(1, 141) = 0.39$ ,  $p = .53$ ,  $\eta_p^2 = .003$ ), nor was there an interaction between condition and gender ( $F(1, 141) = 1.31$ ,  $p = .25$ ,  $\eta_p^2 = .01$ ), demonstrating that both females and males experienced more SNS arousal during the ST than control task. However, as reported in the main text, females in the ST condition perceived more demands relative to resources when compared to males in the ST condition.

#### Association between Math Identification and Demand/Resource Index

For readers interested in understanding the relationship between math identification and the demand/resource index, we predicted the demands/resource ratio from participants' math identification, gender, and condition. We interpret these results with caution given that we intentionally restricted the range on math identification (to be 5 or greater on 1 to 7 scale) in order to study the experiences of women highly identified with math. That being said, we found

a significant main effect of math identification,  $b = -0.10$ ,  $SE = 0.04$ ,  $p = .020$ , such that students more highly identified with math perceived lower demands relative to resources than those who were less identified with math. We did not find any interactions between math identification and gender ( $p = .21$ ), math identification and condition ( $p = .38$ ), or math identification, gender, and condition ( $p = .23$ ).

### **Studies 1A and 1B – Additional Analytic Details**

#### **Physiological Linkage**

In the physiological linkage models, we specified the following random effects: variance in the intercepts, variance in the slopes for respondent PEP reactivity, between-person covariance of Person 1 (P1) intercept with Person 2 (P2) intercept, between-person covariance of P1 slope for respondent PEP and P2 slope for respondent PEP, within-person covariance of intercepts with slopes for respondent PEP, between-person covariance of intercepts with slopes for respondent PEP, and variance of slopes for partner PEP.

#### **Performance**

In Studies 1A and 1B, along with the main effects of condition, time, and an interaction of condition by time, we also included difficulty and whether the question was answered the first or second time (i.e., after the participant had worked alone to solve the problem or after the participant had discussed the problem with his/her partner; referred to as “being together” hereafter; coded as -1 when alone and 1 when together). We also included difficulty by time and being together by time interaction terms (see Fitzmaurice, Laird, & Ware, 2011).

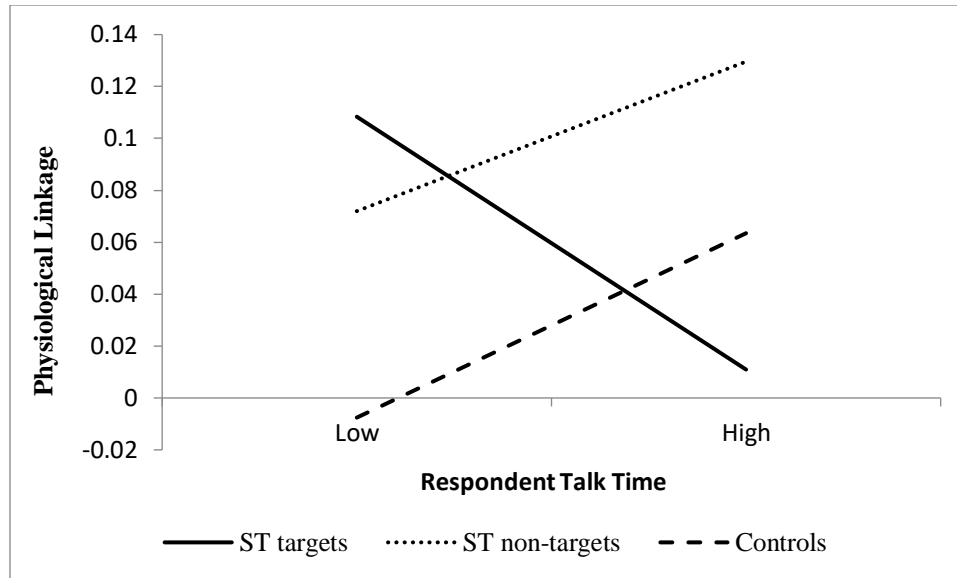
## Study 1A – Additional Results

### SNS Reactivity during the Manipulation

Participants in the ST condition showed significantly greater SNS arousal during the backwards counting task ( $M = -17.74$ ,  $SD = 11.71$ ) than those in the control condition ( $M = -4.23$ ,  $SD = 9.81$ ;  $F(1, 80) = 31.85$ ,  $p < .001$ ;  $\eta_p^2 = .29$ ), including BMI as a covariate (Carroll et al., 2008).

### Physiological Linkage

We observed a significant Partner Prior PEP  $\times$  Condition  $\times$  Respondent Talk Time interaction,  $F(2, 1371) = 5.43$ ,  $p = .005$ , indicating that the strength of the linkage paths were different by condition and further differed by how much one talked about math. A significant two-way Partner Prior PEP  $\times$  Respondent Talk Time interaction was found for ST targets ( $b = -0.01$ ,  $SE = 0.002$ ,  $t(808) = -2.26$ ,  $p = 0.02$ , 95% CI: -0.01 to -0.007) and for controls ( $b = 0.004$ ,  $SE = 0.002$ ,  $t(1367) = 2.40$ ,  $p = 0.016$ , 95% CI: 0.0007 to 0.007) but not for ST non-targets ( $p = .30$ ). For ST targets, the more they talked, the less they were linked to their partners. For controls, the more they talked, the more they were linked to their partners (see Figure S1).



*Figure S1.* Study 1A: Physiological linkage as a function of condition and respondent talk time about math. Predicted values are presented at +/- 1 standard deviation from the mean for partner talk time about math.

## Performance

Recall that we did not find a significant two-way Condition  $\times$  Time interaction, Wald  $\chi^2(2) = 5.11, p = .078$ . However, we report the individual trajectories for each condition for interested readers here. Follow-up analyses revealed that for ST targets and controls, performance improved over time (ST targets:  $b = .01, SE = 0.003, \text{Wald } \chi^2(1) = 3.07, p = .08, 95\% \text{ CI: } -0.001 \text{ to } 0.012$ ; controls:  $b = 0.01, SE = 0.004, \text{Wald } \chi^2(1) = 5.36, p = .021, 95\% \text{ CI: } 0.001 \text{ to } 0.017$ ; see Figure 5). For ST non-targets, performance did not change over time,  $b = -0.001, SE = 0.003, \text{Wald } \chi^2(1) = 0.026, p = .87, 95\% \text{ CI: } -0.01 \text{ to } 0.01$ .

There was a main effect of difficulty (Wald  $\chi^2(2) = 420.53, p < .001$ ), which was qualified by a difficulty by time interaction (Wald  $\chi^2(2) = 38.65, p < .001$ ). Over time, performance got worse on the easy questions ( $b = -0.02, SE = 0.005, \text{Wald } \chi^2(1) = 10.26, p = .001, 95\% \text{ CI: } -0.02 \text{ to } -0.01$ ), better on the medium questions ( $b = 0.02, SE = 0.003, \text{Wald } \chi^2(1) = 35.38, p < .001, 95\% \text{ CI: } 0.01 \text{ to } 0.03$ ), and marginally better on the hard questions ( $b = 0.01, SE = 0.004, \text{Wald } \chi^2(1) = 3.62, p = .057, 95\% \text{ CI: } <0.001 \text{ to } 0.02$ ).

There was a main effect of being together; questions answered after working with one's partner were more likely to be answered correctly than those that were answered after working alone, ( $b = 0.39, SE = 0.05, \text{Wald } \chi^2(1) = 72.94, p < .001, 95\% \text{ CI: } 0.30 \text{ to } 0.48$ ). This effect did not change over time, (Wald  $\chi^2(1) = 0.83, p = .36$ ).

### **Study 1B – Additional Results**

#### **SNS Reactivity during the Manipulation**

ST targets showed significantly greater SNS arousal during the backwards counting task ( $M = -15.41, SD = 12.10$ ) than those who completed the control task ( $M = -4.45, SD = 7.22; F(1, 108) = 29.24, p < .001; \eta_p^2 = .21$ ), including BMI as a covariate.

#### **Performance**

There was a main effect of difficulty (Wald  $\chi^2(2) = 597.29, p < .001$ ), which was qualified by a difficulty by time interaction (Wald  $\chi^2(2) = 55.66, p < .001$ ). Over time, performance got worse on the easy questions ( $b = -.02, SE = 0.005, \text{Wald } \chi^2(1) = 9.53, p < .002, 95\% \text{ CI: } -0.02 \text{ to } -0.01$ ), better on the medium questions ( $b = 0.02, SE = 0.003, \text{Wald } \chi^2(1) = 54.51, p < .001, 95\% \text{ CI: } 0.02 \text{ to } 0.03$ ), and did not change on the hard questions ( $b = -0.002, SE = 0.003, \text{Wald } \chi^2(1) = 0.33, p = .56, 95\% \text{ CI: } -0.01 \text{ to } 0.01$ ).

There was a main effect of being together; questions answered after working with one's partner were more likely to be answered correctly than those that were answered after working alone, ( $b = 0.46$ ,  $SE = 0.03$ , Wald  $\chi^2(1) = 188.44$ ,  $p < .001$ , 95% CI: 0.40 to 0.53). This effect did not change over time, (Wald  $\chi^2(1) = .001$ ,  $p = .98$ ).

## SM References

- Carroll, D., Phillips, A. C., & Der, G. (2008). Body mass index, abdominal adiposity, obesity, and cardiovascular reactions to psychological stress in a large community sample. *Psychosomatic Medicine*, 70(6), 653-660. doi:10.1097/PSY.0b013e31817b9382
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3), 355-391. doi:10.1037/0033-2909.130.3.355
- Fitzmaurice, G. M., Laird, N. M., Ware, J. H. (2011). *Applied longitudinal analysis* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.