Supplemental Methods

Participants

Sixty-one percent of mothers were European American, 20% were Asian, 7% were Latina, and 12% were other or multi-ethnic. There was a range of family income with 13% of participants earning less than $25,000 a year, 32% earning between $26,000 and $75,000 per year, 39.5% earning between $76,000 and $150,000 per year, and 15% earning over $150,000 per year. Seventy-four percent of mothers had earned a bachelor's degree or higher and 96% of mothers were married or living with their partner.

Procedures

During recruitment we instructed mothers to arrive at the lab with an alternate caregiver who was familiar to the baby. The majority of mothers did so (61%), and in the cases where the mother was unable to bring someone, we arranged to have a trained caregiver at the laboratory.

For the 2-minute period during which the infant's baseline cardiovascular responses were obtained, an audio recording of a common lullaby accompanied by a video of floating bubbles was played to help the infant relax and stay still. While not a true baseline due to the infant's possible engagement with the lullaby, this technique is common in developmental psychophysiology research (Alkon et al., 2006) to help reduce artifact in the physiological data due to movement.

For the post-stress interviews with the mother-infant dyad, one of the female interviewers was the evaluator from the TSST or the female experimenter from the control condition (familiar interviewer and the other was completely unfamiliar to the mother (both interviewers were unfamiliar to the infant). The order in which the interviewers conducted the post-stress interviews was counterbalanced. The two female interviewers were matched on general characteristics such as height, build, hair- and eye-color. The interviewers were also trained to have the same general demeanor and affective reactions (i.e., pleasant) towards the mother-infant dyad. We did not observe any differences in behavior or physiology based on whether the mother-infant dyad interacted with the familiar or unfamiliar (to mother)

interviewer so we did not mention this variable again. We may have inadvertently used interviewers too similar in appearance to observe behavioral or physiological differences based on interviewer familiarity.

Supplemental Results

Manipulation check

We examined changes in affect by condition and observed significant effects for both positive and negative affect, $F(2, 64) = 7.18$, $p = .002$ and $F(2, 64) = 3.95$, $p = .02$, respectively. Controlling for baseline affect, mothers who received negative-evaluation showed greater decreases in positive affect ($M = -3.83$, $SD = 5.04$) and little change in negative affect ($M = -.69$, $SD = 6.66$) relative to those receiving positive-evaluation ($M = 2.19$, $SD = 6.28$; $M = -5.36$, $SD = 5.98$), $t(64) = -3.47$, $p = .001$ and $t(64) = 2.80$, $p = .007$, respectively. Those in the control condition showed less decrease in positive affect ($M = 1.30$, $SD = 5.89$) relative to those receiving negative-evaluation, $t(64) = -3.03$, $p = .004$.

Additional analyses

We used BMI as a covariate in all analyses with mothers' physiological reactivity. As expected, BMI was always a significant covariate. For example, in the initial analyses of VC reactivity during the Q&A portion of the stress task, the greater the BMI the smaller the SNS increase, $F(1, 64)= -4.48$, $p=.04$.

We examined whether the person who cared for the infant during the period in which the mother was undergoing the stress manipulation (a familiar or a trained caregiver) was associated with infants' HR reactivity or behavioral avoidance and observed no significant effects, $F(1, 61)=.10$, $p=.75$; $F(1, 63)=.18$, $p=.67$, respectively. We also examined whether infant sex was associated with infants' HR reactivity or behavioral avoidance and again did not observe any significant differences, $F(1, 61)<.01$, $p=.96$; $F(1, 63)=.08$, $p=.77$.

We ran the same analyses that are reported in the main text, selecting out the dyads with missing infant behavioral data. Results are entirely consistent with what is reported in the main text. ANCOVA revealed a significant main effect of mothers' VC reactivity by evaluation condition, $F(2, 61)=6.79$, $p=.002$, with greater VC reactivity for mothers in the positive-evaluation and negative-evaluation compared to the control condition. ANOVA revealed a significant main effect of infants' HR reactivity by

evaluation condition, $F(2,55)=4.83$, $p=.01$, with greater HR reactivity for infants whose mothers received negative-evaluation compared to those whose mothers were in the positive-evaluation or control conditions.

For the covariation model, in which mothers' VC reactivity was treated as the criterion, there was significant variance in the intercept (estimate= 39.44, $SE=7.61$), Wald Z=5.18, $p<.001$, and the linear slope (estimate=.06, $SE=.02$, Wald Z=3.34, $p<.01$). The intercept-slope covariance was positive but not significant (estimate=.21, $SE=.16$, Wald Z=1.25, $p=.21$).