Supplementary Materials: Methods

Participants. Participants (N = 75; 39/36 female/male) ages 18-35 (mean = 24.4; SEM ± 0.4) were recruited from the Linköping University campus via flyers and advertisements. A total of 423 (CC = 251, AC = 147, AA = 25) attended a screening visit to provide a blood sample for prospective genotyping and underwent a psychiatric screening using the Swedish version of the Mini International Neuropsychiatric Interview (M.I.N.I). Exclusion criteria included a lifetime diagnosis of psychosis or bipolar disease, current axis 1 diagnosis, ongoing (within the last month) psychiatric medication, or current (within the last month) use of illegal drugs. Due to the nature of the tasks, anyone with uncorrected visual or hearing impairments was also excluded.

Qualifying participants (N = 25/group; 13 female, 12 male) were invited back to complete the 2-session laboratory study. Participants and research personnel responsible for running behavioral sessions and analyzing data were blinded to participant genotype until after study completion. Participants were paid 750 SEK (approximately 75 Euros) for their participation. All participants provided informed consent prior to participations and the study protocol was approved by the Regional Ethical Review Board, Linköping. Prior to study sessions, participants were asked to refrain from caffeine or tobacco use for 2hr, and eat a standard meal 1hr prior to arrival. Study personnel responsible for running behavioral sessions, scoring psychophysiology data, or analyzing biochemical data did so blinded to genotyping results.

Psychophysiology. Facial EMG sensors consisted of 4mm silver/silver chloride electrodes filled with electrode gel; two placed on each muscle location to form bipolar recording pairs.

Recording electrodes were placed on muscles on the left side of the face and an 8mm ground electrode was placed on the forehead near the hairline. Sites were cleaned with alcohol and lightly abraded and any site with impedance over 20kΩ (measured with a Model 1089 MK III Checktrode; UFI, Morro Bay, CA, USA) was reapplied. EMG signals were amplified, filtered through a 10-500 Hz band pass and 50 Hz comb band stop filter, digitized at 1 kHz, re-filtered, rectified, and integrated over 20ms using EMG100C amplifiers, MP150 Data Acquisition system, and Acqknowledge software from Biopac Systems (Biopac Systems, Inc, Camino Goleta, CA, USA). In all tasks, trials with excessive baseline activity or artefactual activations were identified and excluded by trained, blinded raters. The number of trials excluded based on these factors ranged from 2.5 to 9.5% across muscle location (zygomatic, corrugator, orbicularis) and task.

Affective Image task. Affective images were selected from the International Affective Picture System (IAPS[1]) and divided into four separate sets matched based on normative ratings of valence and arousal. Thus, each participant completed all four sets (pre-stress, post-stress, pre-control, post-control), pseudorandomized. During the task, participants viewed a single image for 6 sec and then rated valence from -4 (negative) to +4 (positive) and physiological arousal on a scale of 0 to 9. Facial EMG recordings of the zygomatic and corrugator were assessed throughout the task. To assess baseline affective responses (e.g. affective responses during the very first affective image task at the first study session), mean EMG amplitude during the 6 sec picture presentation was compared to the immediately preceding 1 sec baseline and averaged across stimulus type; positive, neutral, and negative images. The effect of stress on non-specific muscle activity (i.e. muscle activity in the absence of a stimulus)

was assessed via averaging the 1 sec baseline prior to stimulus onset throughout the entirety of the task, comparing before ('pre') to after ('post') stress and control tasks. Finally, the effect of stress on response to affective stimuli was assessed via EMG activity in response to each stimulus type (positive, neutral, negative). Change scores were calculated from pre to post stress and control tasks.

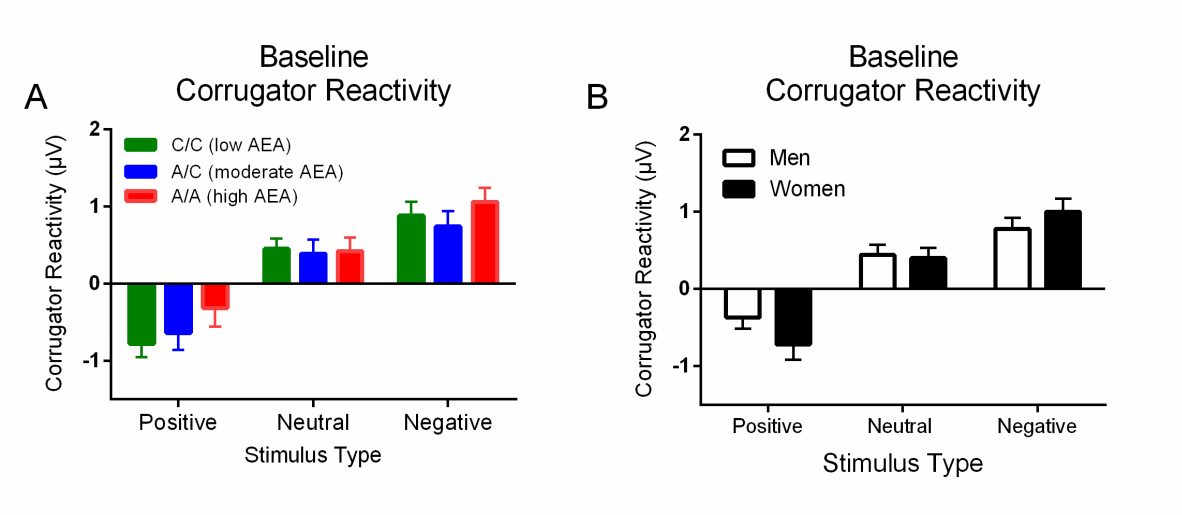
Stress Task. The MAST is a 10 min task consisting of alternating "hand immersion" (HI) trials and "mental arithmetic" (MA) trials[2]. In HI trials, participants were asked to place their left hand into cold water (1-4 degrees Celsius) for up to 90 sec. In MA, participants were asked to perform mental math aloud, with mistakes resulting in negative feedback (e.g. "start over") and successful attempts resulting in prompting to increase speed. In the control version of the MAST, HI trials were similar but the water was kept at room temperature (20-24 degrees Celsius) and MA trials simply required participants to count from 1 to 25 at their own pace.

Statistical analysis. Emotional reactivity at baseline (i.e. during the first exposure to emotional images, prior to stress or control tasks) was assessed using a repeated measures analysis of variance (RM-ANOVA) with the within subject factor of stimulus type (positive, neutral, negative), and genotype and/or gender as between-subjects factors. To assess the effect of stress on resting corrugator muscle activity (e.g. activity in the absence of any stimulus), we calculated the within session change (Post – Pre) in average baseline EMG (in the absence of any emotional stimuli) during at each session and then calculated the effect of stress as the [change at stress session] – [change at control session]. The effect of stress on resting/non-specific corrugator activity was assessed using a one-way ANOVA. Similarly, we calculated the

change (Post – Pre) in EMG value at each session (stress, control) for each stimulus type (positive, neutral, negative). This produced 3 variables representing the effect of stress on i) positive, ii) neutral, and iii) negative stimuli. These variables were analyzed using a RM-ANOVA with stimulus type as the within-subject factor. Thus, these variables represented the overall *change* in reactivity due to stress, not the absolute value of EMG activity. The original analysis was carried out with the between-subjects factor of genotype, as reported in[3]. Here, we ran the same analysis with the between-subjective factor of gender, as well as an additional analysis with both genotype and gender included. For all analyses, significance was set at $p < 0.05$ and followed up with *Bonferroni-Holm* corrected post-hoc tests when appropriate.

Supplementary Materials: Results

There was no effect of genotype or gender on corrugator reactivity during the first exposure to affective images (e.g. "baseline" emotional reactivity). With the between-subjects factor of genotype, we found a main effect of stimulus type ($F(2,138) = 56.1$, $p < 0.001$), btu no main effect of genotype ($p = 0.30$) or type*genotype interaction ($p = 0.48$). Similarly, when gender was the between-subjects factor, we found a main effect of stimulus type ($F(2,140) = 58.4$, $p < 0.001$), but no effect of gender ($p = 0.73$), or type*gender interaction ($p = 0.14$). Finally, when both gender and genotype were entered as between-subjects factors, we again found a main effect of stimulus type ($F(2,132) = 55.9$, $p < 0.001$), but no no effect of gender ($p = 0.64$), genotype ($p = 0.39$), or any other interaction.



Supplemental Figure 1: Baseline corrugator reactivity in response to emotional images. *At the first exposure to the affective image task, there was no difference in corrugator reactivity to positive, neutral, or negative images based either on genotype (A) or gender (B). However, all individuals show the expected main effect of stimulus type, such that negative images elicit the greatest increase in corrugator reactivity ("more frowning") and positive images elicit a reduction in corrugator reactivity ("less frowning").*