PNAS
Proceedings of the National Academy of Sciences of the United States of America

# Supporting Information Appendix for

## Hierarchy Stability Moderates the Impact of Status on Stress and Performance in Humans

Erik L. Knight, Pranjal H. Mehta*

*Corresponding author: mehta@uoregon.edu

**This PDF file includes:**

Data and Materials
Supplementary Analyses
Tables: S1 – S12
Figures: S1 – S4

**Open Data and Materials**

All data and materials for this publication – including experimental manipulation materials, questionnaires administered, and statistical analysis syntax – can be found online (https://osf.io/js2x4).

| | *Full Sample* | *Males* | *Females* |
|---|---|---|---|
| | **Cortisol (µg/dL)** | | |
| Baseline | 0.244 (.020) | 0.232 (.022) | 0.252 (.031) |
| TSST +0’ | 0.322 (.023) | 0.333 (.032) | 0.313 (.033) |
| TSST +20’ | 0.342 (.028) | 0.39 (.043) | 0.307 (.037) |
| TSST +40’ | 0.284 (.024) | 0.313 (.039) | 0.262 (.031) |
| Reactivity[a] | 0.243 (.039) | 0.32 (.064) | 0.185 (.046) |
| Recovery[b] | 0.058 (.054) | 0.19 (.081) | -0.041 (.071) |
| | **Testosterone (pg/ml)** | | |
| Baseline | 112.8 (6.19) | 171.3 (8.36) | 69.1 (2.64) |
| TSST+0’ | 129.0 (7.05) | 195.3 (9.48) | 79.6 (3.28) |
| TSST+20’ | 124.0 (7.18) | 192.0 (9.65) | 73.3 (3.07) |
| TSST+40’ | 113.1 (6.45) | 175.9 (8.20) | 66.2 (2.75) |
| Reactivity[a] | 0.13 (.015) | 0.128 (.023) | 0.132 (.021) |
| Recovery[b] | -0.023 (.022) | 0.019 (.028) | -0.054 (.033) |
| | **Pre-stress Self-report** | | |
| In Control | 2.37 (.10) | 2.77 (.16) | 2.08 (.11) |
| Positive Affect | 2.27 (.08) | 2.57 (.11) | 2.04 (.09) |
| Negative Affect | 2.06 (.07) | 2.02 (.11) | 2.09 (.08) |
| | **Post-stress Self-report** | | |
| In Control | 2.25 (.10) | 2.78 (.17) | 1.85 (.11) |
| Positive Affect | 2.04 (.07) | 2.33 (.11) | 1.83 (.09) |
| Negative Affect | 2.09 (.08) | 2.05 (.12) | 2.12 (.10) |
| | **Behavior During Stressor** | | |
| Interview Performance | 5.15 (.08) | 5.00 (.14) | 5.26 (.09) |
| Dominance | 4.71 (.07) | 4.77 (.11) | 4.66 (.10) |
| Warmth | 4.61 (.08) | 4.40 (.14) | 4.77 (.09) |

**Table S1**. Mean (SE) for study variables in the full study sample and separately for males and females. [a]Reactivity is log-transformed hormone concentrations immediately after the social-evaluative stressor minus baseline concentrations (i.e., log[TSST+0] – log[baseline]). [b]Recovery is log-transformed hormone concentrations forty minutes after the social-evaluative stressor minus baseline concentrations (i.e., log[TSST+40] – log[baseline]).

## Supporting Materials and Methods

### Participants

The ethnic breakdown of participants in the study was approximately 70% European-American, 13% Asian or Asian-American, 7% Hispanic/Latino, 5% Pacific Islander, and 3% or less African-American, Middle Eastern, and Native American. Sample size was estimated *a priori* via power analysis with G*Power3 (1), which assumed four groups with four repeated measures, power ß= 0.80, a small effect size F = 0.15, α = .05, correlation among repeated measures = 0.65, and non-sphericity correction = 0.75.

### Procedure

**Pre-Lab and Arrival.** Prior to arriving at the laboratory session, participants responded to personality questionnaires online, which were used as part of the status manipulation. Participants were instructed to abstain from eating, drinking, exercising, and smoking for two hours before their scheduled experimental session. To account for diurnal variability in endocrine activity, all sessions occurred in the afternoon between 1300 and 1730 hrs. After arriving at the laboratory, participants were seated in an individual testing room where informed consent was obtained to participate in a group activity and perform a speech task. Saliva-sampling and demographic questionnaires were administered for approximately 10 minutes before baseline saliva was collected via passive drool.

**Status Manipulation**. Participants were then told that based on their responses to pre-laboratory questionnaires, they had been assigned to complete an upcoming puzzle task as either the “manager” (high status) or “builder” (low status) while another participant (actually a confederate) would perform the unassigned role. Participants were told specifically that the assignment was based on their “leadership skills and experience” in order to connect this manipulation to prestige (expertise, skills). In actuality, status was randomly assigned and there

| | *Stable Hierarchy* | | *Unstable Hierarchy* | |
|---|---|---|---|---|
| | *High Status* | *Low Status* | *High Status* | *Low Status* |
| | | **Cortisol** | | |
| Reactivity | .10 (.07) | .24 (.07) | .42 (.09) | .23 (.07) |
| Recovery | -.19 (.09) | .11 (.10) | .29 (.12) | .05 (.11) |
| | | **Testosterone** | | |
| Reactivity | .07 (.03) | .14 (.03) | .20 (.03) | .11 (.03) |
| Recovery | -.09 (.04) | -.02 (.04) | .09 (.04) | -.05 (.04) |
| | | **Affect and Behavior** | | |
| Feeling in control | 2.76 (.17) | 2.02 (.16) | 2.17 (.17) | 2.27 (.21) |
| Interview Performance | 5.40 (.14) | 4.72 (.16) | 5.26 (.14) | 5.23 (.15) |
| Dominance | 5.10 (.12) | 4.26 (.13) | 4.79 (.15) | 4.69 (.14) |
| Warmth | 4.83 (.15) | 4.28 (.18) | 4.49 (.19) | 4.72 (.14) |

**Table S2.** Conditional means (SEs) for main dependent variables.

was no puzzle task. All participants were told that the participant in the role of manager would be in charge of directing subordinates, would decide how to structure the process for building the tasks, and would evaluate the "builder" at the end of the task in order to determine how to split $10 of bonus money.

**Stability Manipulation and Social-Evaluative Stressor**. Next, all participants were asked to complete a "speech task in front of a panel of observers" who were "trained in behavioral observation and social competency" in order to "see how [the participant] interact[s] with others." This task is actually the TSST, a well-validated social-evaluative stressor that involves delivering a five-minute speech about one's qualifications for one's ideal job and doing five minutes of serial subtraction in front of two evaluators. The panel of evaluators consisted of a college-aged man and woman (i.e., approximately the same age range as the participants) who were trained to maintain neutral facial expressions and generally be non-reactive. Participants were told that their role (manager/builder) could change based on the speech/math task (unstable hierarchy) or that their performance on the task will not affect their role assignment (stable hierarchy). A five-minute preparation period (but not the speaking portion) was completed in the presence of a gender-matched confederate in order to increase the salience of the manipulations. Panelists and confederates were blind to participants' randomly assigned conditions.

Following completion of the TSST, the participants then recovered for forty minutes while filling out additional demographic questionnaires and performing unrelated tasks not included in the present report. Subsequent saliva samples were collected at 0, 20, and 40 minutes post-TSST for a total of four saliva samples, including baseline.

**Affective States.** After assignment to status and stability conditions and after the TSST, participants responded to a prompt asking how "in control" they felt, which was included as a separate item in a broader measure of self-reported affect. This item was analyzed separately using GLMs because theory suggests that status and hierarchy stability may influence feeling in control specifically, but not necessarily influence general positive or negative affect (2).

**Saliva Sampling and Assays.** In order to collect saliva, participants were instructed to drool approximately 2 mL of saliva into plastic centrifuge tubes, which was immediately frozen in a -20 °C freezer and then transported to a -80 °C freezer for long-term storage. Consistent with standard published procedures (3), saliva samples were later thawed and centrifuged at 3500 rpm for 10 minutes at room temperature. The remaining fluid was then aliquoted into 250 µL samples and frozen again before being thawed and analyzed for cortisol and testosterone in duplicate using enzyme immunoassay kits (Salimetrics, LLC; State College, PA). The average intra-assay coefficients of variation (CVs) were 5.59% (cortisol) and 6.00% (testosterone); the inter-assay CVs were 8.22% (cortisol) and 8.10% (testosterone) averaged across low and high control samples.

**Behavioral Ratings.** Three trained research assistants (2 female), who were naïve to each participant's experimental condition and the purpose of the study, watched the first 2.5 minutes of each participant's speech. They then rated how much they agreed that twenty-nine variables were present in the video, on a scale from 1 – extremely disagree, to 8 – extremely agree. These variables were inspired by previous theory and research on behavioral responses to status and stress (4-9) and represented behavioral components of each participant's competence (e.g., intelligent, competent, etc.), dominance (e.g., confident, dominant, etc.), and warmth (e.g., warm, friendly, etc.) – three theorized behavioral routes to status attainment (8,10). The research assistants answered two additional questions regarding the participant's overall interview performance on separate scales: "How good was this interview?" (1 – extremely bad to 8 – extremely good) and "If you were in charge of hiring, how likely would you be to hire this individual?" (1 – extremely unlikely, to 8 – extremely likely). In order to reduce the potential for gender stereotypes to influence ratings, all male participants' videos were watched and rated in random order before female participants' videos were watched and rated in random order.

The research assistants' responses (average inter-rater reliability across all variables: $\alpha = .665$) were submitted to a factor analysis. A three-factor solution with varimax rotation was investigated and found to account for 66.4% of variance (see Table S9).

**Statistical Analyses**

**Data Transformation.** Two-tailed Kolmogorov-Smirnov tests of normality revealed non-normal distributions for cortisol and testosterone concentrations at multiple time points (*p*s < .03). We corrected this non-normality by natural-log-transforming cortisol and testosterone concentrations and used these transformed scores in analyses that examined the effects of the experimental manipulations on changes in endocrine concentrations over time. The scale that cortisol is measured on (e.g., Baseline concentration, M = .244 μg/dL, SE = .020, range = [.07, 1.84]) results in negative values for many of the log-transformed cortisol data. Thus, an arbitrary value of 10 was added to each transformed cortisol value to ensure that all values were positive for ease of interpretation. Testosterone's scale (e.g., Baseline concentration, M = 112.8, SE = 6.19, range = [30.4, 343.2]) does not result in negative transformed values and so did not require an arbitrary linear transformation.

**Endocrine Analyses.** To analyze the overall endocrine response patterns, 2 (High vs. Low Status) x 2 (Stable vs. Unstable Hierarchy) x 4 (Time) repeated measures GLM were used. Mauchly's test of sphericity revealed violations of sphericity, so Huynh-Feldt corrections were applied. For follow-up analyses of acute reactivity, each hormone's change from baseline to immediately after the stressor (TSST+0) was calculated. Similarly for recovery, endocrine change from baseline to forty minutes after the stressor was calculated. This index of recovery measures the extent to which individuals were exposed to a given hormone during a period in which hormones should decline following initial reactivity to the stressor (11). Larger, positive values indicate a hormone did not return to baseline during the forty minutes of recovery; a zero or negative value indicates a hormone did return to baseline (or sub-baseline levels consistent with circadian decline in hormone concentrations). These values were regressed on status, hierarchical stability, and their interaction (in addition to participant sex for testosterone analyses) in separate univariate GLMs.

**Behavioral Analyses.** We conducted separate GLM analyses on interview performance, dominance, and warmth with status, stability, and their interaction as between subject variables with participant sex as a covariate.

**Moderated Mediation Analyses.** Using the *PROCESS* Model 8 template in *SPSS* (v. 22, IBM Corp.), our primary moderated mediation models were produced with interview performance as the outcome variable; social status as the independent variable; feeling in control, testosterone reactivity, or cortisol reactivity as the mediator; sex as a covariate; and hierarchy stability as a moderator. We also produced similar moderated mediation models with dominance or warmth as outcome variables. Bootstrap analyses were used to calculate bias-corrected 95%

confidence intervals for the indirect effects of each putative mediator (n = 1000 subsamples). Indices of reactivity were based on standardized residuals that were produced from regressing TSST+0 concentrations on baseline concentrations. Residuals for recovery were calculated from regressing TSST+40 endocrine concentrations on baseline endocrine concentrations. For testosterone, these residual values were normally distributed when produced from raw testosterone concentrations. Raw cortisol concentrations resulted in skewed residuals, and so log-transformed cortisol concentrations were submitted to residual calculation and used in the moderated mediation models. These metrics of endocrine reactivity and recovery were employed in the correlational analyses (Table S11) and in the moderated mediation analyses.

**Supplementary Analyses**

In the sections below, we report supplemental analyses for (a) endocrine reactivity and recovery; (b) the moderating effects of sex; (c) robustness checks for the main analyses; (d) behaviors during the stressor; (e) moderated mediation analyses; and (f) positive and negative affect.

**Supplementary Endocrine Analyses**

**Alternative Methods for Measuring Acute Endocrine Reactivity.** The main text reports analyses for hormone reactivity to the stressor as change scores from Baseline to TSST + 0. We confirmed that the Status x Stability effects on endocrine reactivity extend to the following alternative methods for modeling acute reactivity: percent change scores; unstandardized residuals calculated by regressing endocrine concentrations at TSST+0 on the baseline concentrations; and area-under-the-curve with respect to increase (AUCi; 12). AUCi was calculated as:

$$AUC_i = \left(\frac{Baseline + TSST40}{2}\right) + TSST0 + TSST20 - (4 - 1) * Baseline$$

Distinguishing it from the other three indices of reactivity, AUCi takes into account all four samples and thus represents change in endocrine concentrations across the reactivity and recovery period. Tables S3 and S4 (and Figures S1 and S2) show results for cortisol and testosterone, respectively, across these different analyses.

| | Raw Change | | | Percent Change | | | Unstandardized Residuals | | | AUCi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | CI | p | B | CI | p | B | CI | p | B | CI | p |
| **(Intercept)** | 0.25 | 0.17 – 0.32 | **<.001** | 0.03 | 0.02 – 0.04 | **<.001** | 0.00 | -0.07 – 0.08 | .907 | 0.52 | 0.29 – 0.75 | **<.001** |
| **Social Status** | 0.01 | -0.06 – 0.09 | .734 | 0.00 | -0.01 – 0.01 | .800 | 0.01 | -0.06 – 0.09 | .725 | 0.02 | -0.21 – 0.24 | .877 |
| **Hierarchy Stability** | -0.08 | -0.15 – -0.00 | **.039** | -0.01 | -0.02 – -0.00 | **.034** | -0.07 | -0.15 – 0.00 | .060 | -0.24 | -0.47 – -0.01 | **.037** |
| **Status x Stability** | -0.08 | -0.16 – -0.01 | **.030** | -0.01 | -0.02 – -0.00 | **.030** | -0.08 | -0.16 – -0.01 | **.025** | -0.25 | -0.48 – -0.02 | **.031** |
| Observations | 110 | | | 110 | | | 110 | | | 110 | | |
| $R^2$ / adj. $R^2$ | .082 / .056 | | | .080 / .054 | | | .077 / .051 | | | .035 / .008 | | |

**Table S3.** Cortisol reactivity modeled in several ways. Raw change was reported in the main text but is reported here as well for comparison purposes. Each column represents a separate GLM with effects coded variables (Status: High = 1, Low = -1; Stability: Stable = 1, Unstable = -1).

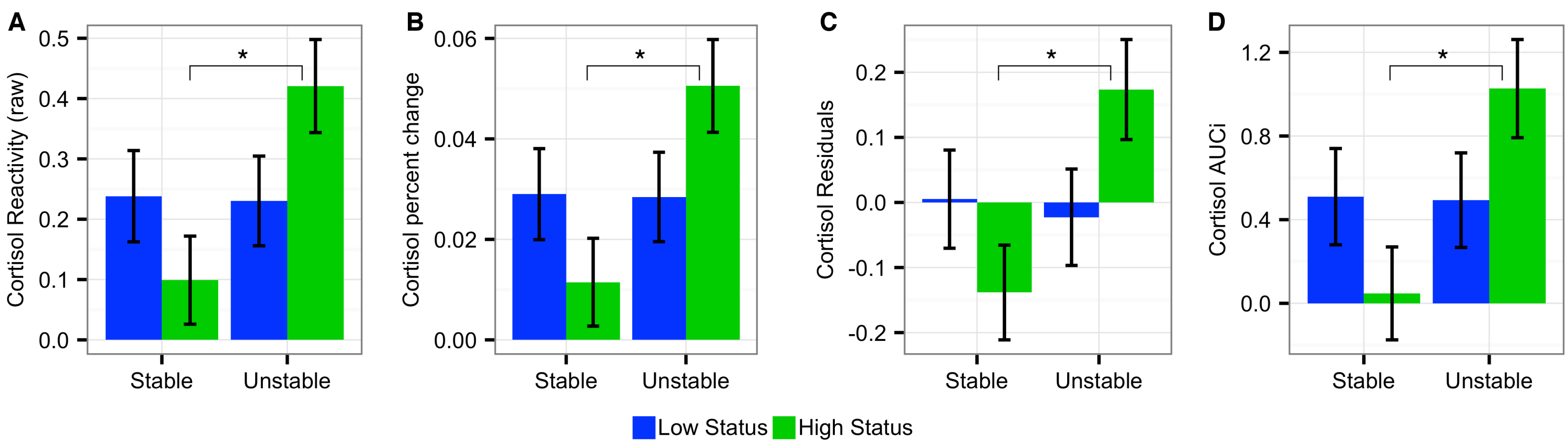


**Figure S1**. Cortisol reactivity modeled in several ways. **A**. Raw difference **B**. Percent change **C**. Residual change **D**. $AUC_i$ * = significant uncorrected pairwise comparison at $p < .05$

| | Raw Change | | | Percent Change | | | Residuals | | | AUCi | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | CI | p | B | CI | p | B | CI | p | B | CI | p |
| **(Intercept)** | 0.13 | 0.10 – 0.16 | **<.001** | 0.03 | 0.02 – 0.04 | **<.001** | 0.00 | -0.03 – 0.03 | .814 | 0.21 | 0.13 – 0.28 | **<.001** |
| **Social Status** | 0.00 | -0.03 – 0.03 | .761 | 0.00 | -0.01 – 0.01 | .751 | 0.01 | -0.02 – 0.03 | .725 | 0.03 | -0.05 – 0.10 | .507 |
| **Hierarchy Stability** | -0.03 | -0.06 – 0.00 | .090 | -0.01 | -0.01 – 0.00 | .101 | -0.02 | -0.05 – 0.01 | .107 | -0.08 | -0.15 – -0.00 | **.048** |
| **Sex** | -0.00 | -0.03 – 0.03 | .987 | -0.00 | -0.01 – 0.00 | .418 | 0.01 | -0.02 – 0.04 | .454 | 0.05 | -0.03 – 0.13 | .195 |
| **Status x Stability** | -0.04 | -0.07 – -0.01 | **.008** | -0.01 | -0.02 – -0.00 | **.011** | -0.04 | -0.07 – -0.01 | **.008** | -0.12 | -0.20 – -0.05 | **.002** |
| Observations | 110 | | | 110 | | | 110 | | | 110 | | |
| $R^2$ / adj. $R^2$ | .090 / .055 | | | .090 / .056 | | | .090 / .055 | | | .131 / .097 | | |

**Table S4.** Testosterone reactivity modeled in several ways. Raw change was reported in the main text but is reported here as well for comparison purposes. Each column represents a separate GLM with effects coded variables (Status: High = 1, Low = -1; Stability: Stable = 1, Unstable = -1; Sex: Male = 1, Female = -1).

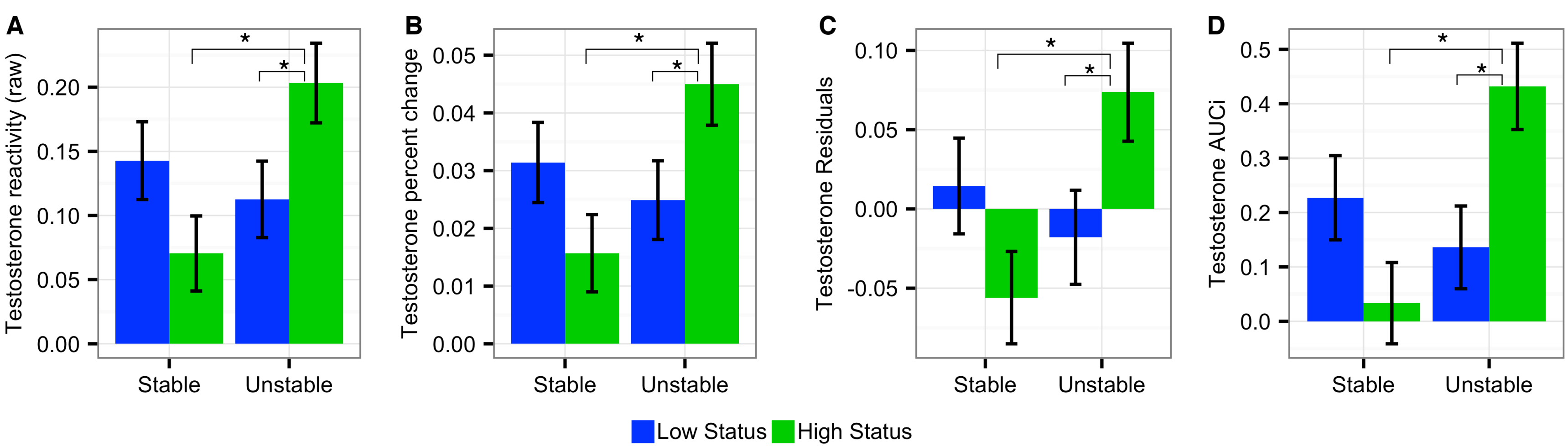


**Figure S2.** Testosterone reactivity modeled in several ways. **A**. Raw difference **B**. Percent change **C**. Residual change **D**. $AUC_i$ * = significant uncorrected pairwise comparison at p < .05

**Endocrine Recovery Slope.** The main text reports hormone recovery as change scores from Baseline to TSST +40. We also examined the effect of status and hierarchy instability on recovery slope. To calculate recovery slope, we used the lme4 package in R (13) to extract Empirical Bayes estimates of the linear slope that connects the three post-stressor samples (TSST+0, TSST+20, TSST+40) for each participant. This recovery slope represents a bias-corrected measurement of the rate at which participants' hormone concentrations changed over the three post-stressor samples and is appropriate for between-person comparisons of endocrine recovery (14, 15). Within this measure, more negative numbers represent a quicker reduction (steeper slope) following activation, and less negative numbers represent a more prolonged recovery period (flatter slope).

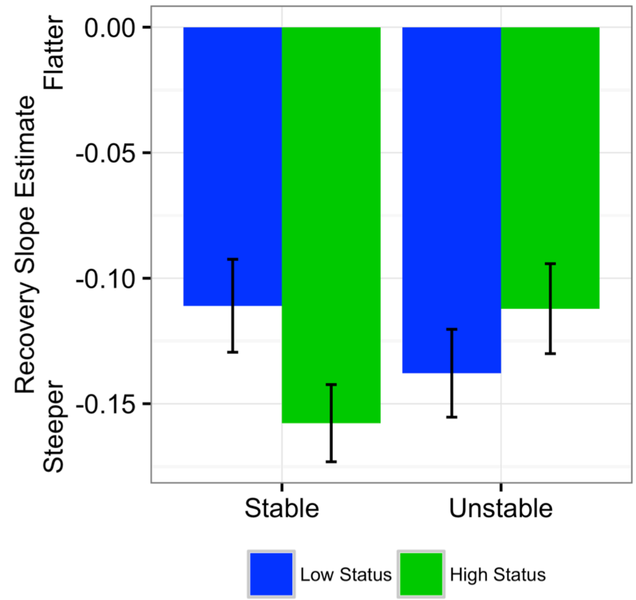


**Figure S3.** Cortisol recovery slope, calculated by extracting linear slope from TSST+0 to TSST+40

A GLM found a significant Status x Stability interaction on the recovery slope for cortisol ($F(1,106) = 4.38$, $p = 0.039$, $\eta^2 = 0.040$; Figure S3) but not testosterone ($F(1,105) = 0.564$, $p = 0.454$, $\eta^2 = 0.005$). The pattern of the interaction for cortisol indicates that high status individuals in a stable hierarchy exhibited steeper slopes, indicative of a quicker recovery following activation of the HPA axis. But in the unstable hierarchy, high status individuals had flatter slopes, indicative of extended activation of the HPA axis. The opposite pattern was observed for low status individuals in stable versus unstable hierarchies, although none of the pairwise comparisons were significant.

**Moderating Effect of Participant Sex**

We explored sex as a moderator of the effects of status and hierarchy stability on all dependent variables reported in the main document (Table S5). Consistent with previous research (16), a Time x Sex effect was found for cortisol wherein men showed stronger reactivity to the stressor compared to women ($F(1.91, 194.57) = 6.28$, $p = 0.003$, $\eta^2 = 0.058$). However, the Time x Sex x Status x Stability interactions were non-significant for cortisol and testosterone ($p$-values $> 0.16$, $\eta^2 < 0.17$), suggesting that endocrine responses to social status and hierarchical instability did not depend on participant sex.

There was a Sex x Status x Stability interaction on dominance behavior (see Table S5). The pattern of the interaction revealed that the Status x Stability interaction on dominance was stronger in men than in women. There were non-significant interactions between status, stability, and sex for all other dependent variables reported in the main text. Collectively, these analyses revealed that the status x stability interactions on our primary dependent variables generally showed similar effects in men and women.

| | Cortisol reactivity | | Testosterone reactivity | | Feeling in control | | Interview Performance | | Dominance | | Warmth | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | P | B | p | B | P | B | P | B | p | B | p |
| (Intercept) | 0.26 | **<.001** | 0.13 | **<.001** | 2.35 | **<.001** | 5.14 | **<.001** | 4.71 | **<.001** | 4.59 | **<.001** |
| Status | 0.01 | .713 | 0.00 | .853 | 0.17 | **.041** | 0.18 | **.016** | 0.23 | **.001** | 0.04 | .624 |
| Hierarchy Stability | -0.08 | **.032** | -0.03 | .068 | 0.08 | .325 | -0.09 | .232 | -0.03 | .713 | 0.00 | .972 |
| Sex | 0.07 | .052 | 0.00 | .945 | 0.38 | **<.001** | -0.12 | .103 | 0.05 | .448 | -0.19 | **.024** |
| Status x Stability | -0.09 | **.022** | -0.04 | **.005** | 0.21 | **.011** | 0.17 | **.023** | 0.20 | **.004** | 0.18 | **.029** |
| Status x Sex | -0.01 | .772 | -0.02 | .172 | 0.11 | .205 | 0.12 | .117 | 0.01 | .923 | -0.11 | .165 |
| Stability x Sex | -0.01 | .894 | -0.01 | .349 | 0.02 | .789 | -0.01 | .856 | 0.05 | .496 | -0.08 | .303 |
| Status x Stability x Sex | -0.03 | .434 | -0.02 | .141 | 0.06 | .503 | 0.11 | .146 | 0.14 | **.047** | 0.03 | .741 |
| Observations | 110 | | 110 | | 111 | | 109 | | 109 | | 109 | |
| $R^2$ / adj. $R^2$ | .123 / .063 | | .125 / .065 | | .260 / .210 | | .166 / .108 | | .202 / .146 | | .117 / .056 | |

**Table S5.** Status x Stability x Sex GLMs showing generally null moderating effects of participant sex. Status, stability, and sex are effects coded (Status: High = 1, Low = -1; Stability: Stable = 1, Unstable = -1; Sex: Male = 1, Female = -1).

**Additional robustness checks**

We conducted two additional robustness checks:

1) We tested the extent to which the endocrine results remained robust when controlling for covariates relevant to endocrine function (participant sex, time since awakening, and hours of sleep prior to the experimental session; Table S6) and socioeconomic status (subjective social status via the "ladder" survey (17); mother's and father's education; and family income; Table S7). These analyses revealed statistically significant status x hierarchy stability interactions across all models.

2) We examined models with bias-corrected bootstrap estimates of the status x stability interaction term for the six main GLMs (endocrine reactivity, sense of control, and the three behavioral factors) using the "boot" library in R (18, 19). The models were replicated 1000 times and the bias-corrected and accelerated (BCa) bootstrap estimates of the 95% confidence intervals were extracted. Effects were considered robust if the 95% confidence intervals for the status x stability interaction term did not contain zero. For each model, the social status x hierarchy stability interaction term was robust to bootstrap bias correction (Table S8).

| | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|
| **A**: ΔCortisol | *B* | *p* | *B* | *p* | *B* | *p* |
| Intercept | 0.25 | **<.001** | 0.26 | **<.001** | 0.44 | .159 |
| Social Status | 0.01 | .741 | 0.01 | .724 | 0.02 | .675 |
| Hierarchy Stability | -0.08 | **.041** | -0.08 | **.034** | -0.08 | **.041** |
| **Status x Stability** | -0.08 | **.031** | -0.09 | **.025** | -0.08 | **.031** |
| Sex | | | 0.07 | .052 | 0.08 | .053 |
| Time since awakening | | | | | -0.01 | .530 |
| Hrs of sleep | | | | | -0.01 | .697 |
| Observations | 109 | | 109 | | 108 | |
| $R^2$ / adj. $R^2$ | .080 / .054 | | .117 / .083 | | .119 / .066 | |

| | Model 1 | | Model 2 | |
|---|---|---|---|---|
| **B**: ΔTestosterone | *B* | *p* | *B* | *p* |
| Intercept | 0.13 | **<.001** | 0.13 | .307 |
| Social Status | 0.00 | .797 | 0.01 | .705 |
| Hierarchy Stability | -0.03 | .102 | -0.03 | .063 |
| Sex | 0.00 | .973 | 0.00 | .909 |
| **Status x Stability** | -0.04 | **.007** | -0.04 | **.008** |
| Time since awakening | | | 0.01 | .449 |
| Hrs of sleep | | | -0.01 | .606 |
| Observations | 109 | | 109 | |
| $R^2$ / adj. $R^2$ | .090/.055 | | .102/.048 | |

**Table S6.** Cortisol and testosterone reactivity to stress controlling for biosocial variables. **A.** Log-transformed cortisol change (TSST+0 – baseline) regressed on the listed variables. **B.** Log-transformed testosterone change regressed on the listed variables. All categorical variables are effects coded (Status: High = 1, Low = -1; Stability: Stable = 1, Unstable = -1; Sex: Male = 1, Female = -1).

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
|---|---|---|---|---|---|---|---|---|
| **A**: ΔCortisol | B | p | B | p | B | p | B | p |
| Intercept | 0.25 | **<.001** | 0.44 | **.006** | 0.18 | .148 | 0.41 | .110 |
| Social Status | 0.01 | .741 | 0.02 | .569 | 0.02 | .687 | 0.03 | .503 |
| Hierarchy Stability | -0.08 | **.041** | -0.08 | **.048** | -0.08 | **.039** | -0.08 | **.043** |
| Status x Stability | -0.08 | **.031** | -0.09 | **.034** | -0.08 | **.033** | -0.09 | **.033** |
| Mother's Education | | | -0.01 | .751 | | | -0.01 | .774 |
| Father's Education | | | -0.01 | .666 | | | -0.01 | .664 |
| Family Income | | | -0.01 | .454 | | | -0.02 | .471 |
| Subjective SES | | | | | 0.02 | .565 | 0.01 | .780 |
| Observations | 109 | | 105 | | 107 | | 103 | |
| $R^2$ / adj. $R^2$ | .080 / .054 | | .102 / .047 | | .085 / .049 | | .107 / .041 | |

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
|---|---|---|---|---|---|---|---|---|
| **B**: ΔTestosterone | B | p | B | p | B | p | B | p |
| Intercept | 0.13 | **<.001** | 0.20 | **.003** | 0.05 | .307 | 0.06 | .559 |
| Social Status | 0.00 | .797 | 0.01 | .675 | 0.00 | .863 | 0.01 | .728 |
| Hierarchy Stability | -0.03 | .102 | -0.03 | .071 | -0.02 | .128 | -0.03 | .076 |
| Sex | 0.00 | .973 | 0.00 | .848 | 0.00 | .752 | 0.00 | .807 |
| Status X Stability | -0.04 | **.007** | -0.04 | **.008** | -0.04 | **.008** | -0.04 | **.010** |
| Mother's Education | | | -0.01 | .486 | | | -0.01 | .568 |
| Father's Education | | | 0.00 | .798 | | | 0.00 | .656 |
| Family Income | | | -0.01 | .451 | | | -0.00 | .929 |
| Subjective SES | | | | | 0.02 | .080 | 0.02 | .114 |
| Observations | 109 | | 105 | | 107 | | 103 | |
| $R^2$ / adj. $R^2$ | .088 / .053 | | .105 / .060 | | .112 / .048 | | .122 / .048 | |

**Table S7.** Cortisol and testosterone reactivity, controlling for relevant socioeconomic status variables. **A.** Log-transformed cortisol change (TSST+0 – baseline) regressed on the listed variables. **B**. Log-transformed testosterone change (TSST+0 – baseline) regressed on the listed variables. All categorical variables are effects coded (Status: High = 1, Low = -1; Stability: Stable = 1, Unstable = -1; Sex: Male = 1, Female = -1).

| | $B_{original}$ | $B_{robust}$ | 95%CI |
|---|---|---|---|
| Cortisol Reactivity | -0.0827 | -0.0832 | (-0.158, -0.007) |
| Testosterone Reactivity | -0.0415 | -0.0418 | (-0.072, -0.013) |
| Sense of Control | 0.209 | 0.210 | (0.04, 0.386) |
| Interview Performance | 0.170 | 0.171 | (0.025, .317) |
| Dominance | 0.193 | 0.189 | (0.067, 0.338) |
| Warmth | 0.169 | 0.170 | (0.015, 0.326) |

**Table S8**. Bias-corrected bootstrap estimates (r = 1000 subsamples) of the status x stability interaction term and its 95% confidence interval (extracted via BCa method) from each analysis. Each row represents a GLM used in the main document to analyze the DV listed in the first column.

**Supplementary Behavioral Analyses**

Table S9 reports factor loadings and inter-rater reliabilities. Bold numbers indicate that the item loaded on a single factor at > 0.5 and therefore was included in that factor[1].

The inter-rater reliabilities for individual items are generally in line with other research on status-relevant behaviors (e.g., 8). We also examined the inter-rater reliabilities for each behavioral factor (bottom row, Table S9). This metric of inter-reliability is appropriate because our statistical analyses employed these aggregated factors (20, 21). Each rater's scores were averaged into interview performance, dominance, and warmth behavioral factors prior to calculating Cronbach's α for inter-rater reliability. Doing so revealed higher inter-rater reliabilities (Interview Performance: $\alpha = 0.835$; Dominance: $\alpha = 0.834$; Warmth: $\alpha = 0.769$), suggesting that the raters generally agreed on the aggregate measures of behavior.

Despite achieving high inter-reliability at the behavioral factor level, the inter-rater reliabilities for some of the individual items indicated low to moderate agreement among raters. Thus, we tested whether the status x stability interaction on behaviors during the stressor would show the same general pattern after excluding items with inter-rater reliabilities of Cronbach's $\alpha < .60$. This cutoff removed six items – five items from the dominance factor (nervous, stressed, awkward, strong posture, and dominant appearance), one from the warmth factor (humorous), and none from the interview performance factor. This subset of items raised the average inter-rater reliability to $\alpha = 0.714$. We then used GLMs to regress the new behavioral factors on social status, hierarchy instability, and their interaction (controlling for participant sex). As shown in Table S10, the interaction between status and stability remained statistically significant for dominance and was marginally significant for warmth, though the effect was in the same direction and magnitude as the original analysis. Overall, these new analyses reveal the same pattern of results as the main analyses.

[1] Observers also rated participants on six additional items that did not satisfactorily on any one of the factors and were excluded from statistical analyses. These items were: Fidgets with hands, bodily motion, etc.; fidgets with items like a pencil, study equipment, etc.; likeable; maintains eye contact; talks fast; and stumbles over words.

| | Interview Performance | Dominance | Warmth | Inter-rater Reliability (Cronbach's α) |
|---|---|---|---|---|
| How good was this interview? | **0.832** | 0.364 | 0.267 | 0.827 |
| Competent | **0.830** | 0.353 | 0.231 | 0.640 |
| Intelligent | **0.825** | 0.180 | 0.212 | 0.635 |
| How likely would you hire person? | **0.824** | 0.360 | 0.285 | 0.808 |
| Engaged | **0.775** | 0.280 | 0.378 | 0.674 |
| Coherent | **0.747** | 0.446 | 0.089 | 0.702 |
| Bored | **-0.742** | -0.055 | -0.427 | 0.612 |
| Creative | **0.546** | 0.312 | 0.337 | 0.668 |
| Leader-like | 0.282 | **0.819** | 0.303 | 0.778 |
| Confident | 0.345 | **0.818** | 0.307 | 0.827 |
| Nervous | -0.185 | **-0.817** | 0.132 | 0.525 |
| Follower-like | -0.108 | **-0.814** | -0.138 | 0.649 |
| Dominant sounding | 0.115 | **0.808** | 0.258 | 0.724 |
| Stressed | -0.147 | **-0.758** | 0.132 | 0.574 |
| In control | 0.556 | **0.735** | 0.140 | 0.795 |
| Awkward | -0.161 | **-0.711** | 0.051 | 0.543 |
| Strong posture | 0.144 | **0.705** | 0.074 | 0.587 |
| Quiet | -0.089 | **-0.686** | -0.517 | 0.727 |
| Dominant appearing | -0.112 | **0.670** | 0.292 | 0.574 |
| High Status | 0.220 | **0.500** | 0.380 | 0.650 |
| Warm | 0.456 | 0.079 | **0.799** | 0.679 |
| Happy | 0.301 | 0.166 | **0.863** | 0.774 |
| Friendly | 0.379 | 0.081 | **0.839** | 0.677 |
| Smile-y (smiled a lot) | 0.147 | 0.053 | **0.810** | 0.716 |
| Humorous | 0.118 | 0.178 | **0.696** | 0.350 |
| Inter-rater Reliability of aggregate behavior (Cronbach's α) | **0.835** | **0.834** | **0.769** | |

**Table S9.** Factor loadings and inter-rater reliabilities for behavioral items, as assessed by independent observers' ratings of the videotaped social-evaluative stressor.

| | Dominance | | | | Warmth | | | |
|---|---|---|---|---|---|---|---|---|
| | Original | | New (reduced) | | Original | | New (reduced) | |
| | B | p | B | p | B | p | B | p |
| (Intercept) | 4.73 | **<.001** | 4.76 | **<.001** | 4.59 | **<.001** | 4.83 | **<.001** |
| Social Status | 0.25 | **.001** | 0.26 | **.001** | 0.05 | .566 | 0.06 | .476 |
| Hierarchy Stability | -0.06 | .421 | -0.07 | .344 | 0.01 | .907 | 0.02 | .868 |
| Sex | 0.07 | .321 | 0.06 | .437 | -0.18 | **.039** | -0.21 | **.023** |
| Status x Stability | 0.19 | **.008** | 0.23 | **.004** | 0.17 | **.041** | 0.16 | .074 |
| Observations | 106 | | 106 | | 106 | | 106 | |
| $R^2$ / adj. $R^2$ | .182 / .149 | | .181 / .148 | | .081 / .044 | | .082 / .045 | |

**Table S10.** Comparison of GLMs for behavior factors with and without items that had lower inter-rater reliability. Interview performance is not displayed because it did not contain items with lower inter-rater reliability. Status, stability, and sex are effects coded (Status: High = 1, Low = -1; Stability: Stable = 1, Unstable = -1; Sex: Male = 1, Female = -1).

**Supplemental Moderated Mediation Analyses**

We report moderated mediation analyses for interview performance in the main document; here we report the partial correlations among the main variables controlling for sex (Table S11), the conditional indirect effects for those analyses (Table S12), as well as moderated mediation results for dominance and warmth. The Status x Stability interaction on dominance was mediated by feeling in control ($\omega = 0.098$, 95%CI [0.028, 0.205]) but not testosterone or cortisol reactivity (95% CI's overlapped with zero). These results extend prior research (2) by showing that hierarchical instability disrupts the effect of high status on dominance behaviors via reduced feelings of control. The Status x Stability interaction on warmth was not significantly mediated by feeling in control, testosterone reactivity, or cortisol reactivity (95% CIs overlapped with zero). Testosterone and cortisol recovery were not found to mediate any of the behaviors (95% CIs overlapped with zero). These non-significant mediations suggest that other psychological and biological factors that were not measured in the present experiment may explain the effects of the hierarchy on warmth (e.g., progesterone changes, which have been linked to affiliation motivation, 22). Additional studies will be required to identify the mechanisms through which the social hierarchy influences warmth behavioral responses to stress.

Our primary correlational and mediation analyses revealed that greater sense of control was positively related to interview performance, whereas testosterone reactivity was negatively related to interview performance. Additional analyses revealed non-significant sex x testosterone reactivity and sex x feeling in control interactions on interview performance (*p*s> .10, $\eta^2$s < .026). These results suggest that the pathways between these mediators and interview performance did not statistically differ between male and female participants.

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* |
|---|---|---|---|---|---|---|---|
| *1.Cortisol Reactivity* | | | | | | | |
| *2. Cortisol Recovery* | 0.75** | | | | | | |
| *3. T Reactivity* | 0.52** | 0.40** | | | | | |
| *4. T Recovery* | 0.20* | 0.43** | 0.38** | | | | |
| *5. In-Control* | -0.09 | -0.13 | $-0.17^{\dagger}$ | 0.03 | | | |
| *6. Interview Performance* | -0.004 | -0.08 | -0.23* | -0.01 | 0.35** | | |
| *7. Dominance* | -0.05 | -0.14 | -0.07 | -0.05 | 0.37** | 0.61** | |
| *8. Warmth* | 0.05 | -0.02 | -0.10 | -0.07 | 0.15 | 0.55** | 0.41** |

**Table S11.** Partial correlations (controlling for sex) between the primary dependent measures. Reactivity and recovery are calculated by regressing endocrine concentrations at TSST+0 or TSST+40 (respectively) on baseline endocrine concentrations. Missing data is deleted listwise, as is the case in the moderated mediation models. **p < 0.001; *p < 0.05; $^{\dagger}$p < 0.10

| | In Control | | Testosterone Reactivity | |
|---|---|---|---|---|
| | Stable | Unstable | Stable | Unstable |
| Interview Performance | .11<br>[.040, .232] | -.004<br>[-.09, .067] | .045<br>[.000, .139] | -.043<br>[-.12, -.004] |
| Dominance | .099<br>[.045, .183] | -.003<br>[-.07, .06] | --- | --- |
| | Low Status | High Status | Low Status | High Status |
| Interview Performance | -.037<br>[-.154, .021] | .077<br>[.022, .188] | -.013<br>[-.081, .013] | .020<br>[.000, .151] |
| Dominance | -.033<br>[-.118, -.015] | .069<br>[.029, .141] | --- | --- |

**Table S12.** The conditional indirect effects (ω) for each significant moderated mediation with bias-corrected bootstrapped 95% confidence intervals. Status and hierarchy stability are effects coded (Status: High = 1, Low = -1; Stability: Stable = 1, Unstable = -1); all analyses include participant sex as a covariate (Males = 1; Females = -1). The top section of the table shows conditional indirect effects with hierarchy stability as the moderator of the influence of status on interview performance and dominance. The bottom section of the table shows the conditional indirect effects with status as the moderator of the influence of hierarchy stability on interview performance and dominance. We include both sets of conditional effects to inform follow-up research.

**Supplemental Analyses for Self-Reported Affect**

Participants responded to thirteen items related to their momentary positive and negative affect on a 1 to 5 scale, from “Not at all” to “Extremely.” These questions were administered after having status and stability assigned and immediately after the stressor. Positive affect: Interested, excited, happy, strong, enthusiastic, proud, and self-confident (Cronbach’s α =0.89). Negative affect: Distressed, upset, sad, irritable, ashamed, and nervous (Cronbach’s α =0.82). The aggregated positive and negative affect scores were submitted to separate 2 (Time) x 2 (High vs. Low Status) x 2 (Stable vs. Unstable Hierarchy) mixed GLM analyses. There was a marginally significant Status x Stability x Time interaction on positive affect ($F(1,104) = 3.50$, $p = 0.064$, $\eta^2 = 0.033$) but not negative affect ($F(1,104) = 0.958$, $p = 0.330$, $\eta^2 = 0.01$). As shown in Figure S4, the pattern of the interaction aligns with the hierarchy instability hypothesis. The stronger effects for feeling in control reported in the main text compared to the results reported here are consistent with social hierarchy theories, which posit that social rank influences behavior through perceived controllability as opposed to global positive or negative affect (2).

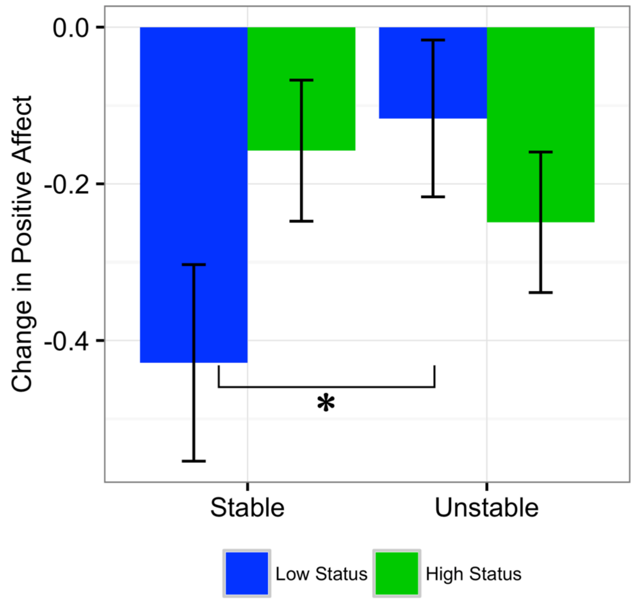


**Figure S4.** Change in self-reported positive affect from pre- to post-stress. Error bars represent standard errors of the mean. * = significant uncorrected pairwise comparison at p < .05

**Additional Acknowledgments**

We also thank research assistants for their assistance in data collection: Isaac Wiggins, James Cragun, Megan Bruun, Patrick Sequeira, Adam Tratner, Angelica Pirumova, Tad Falk, Stefanie Stewart, Jacqueline Trunnel, Nikki Gonzalez, Saejoong Oh, and Tallon Lamoreaux.

## References


1. Faul F, Erdfelder E, Buchner A, & Lang, AG (2009) Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160
2. Fast NJ, Gruenfeld DH, Sivanathan N, & Galinsky AD (2009) Illusory control a generative force behind power's far-reaching effects. *Psychol Sci*, 20(4): 502-508.
3. Schultheiss OC & Stanton SJ (2009) Assessment of salivary hormones. *Methods in social neuroscience*, eds Harmon-Jones E, Beer J(Guilford Publications, New York), pp 17-44.
4. Lammers J, Dubois D, Rucker DD, & Galinsky AD (2013) Power gets the job: priming power improves interview outcomes. *J Exp Soc Psychol*, 49(4): 776-779.
5. Schmid PC, & Schmid Mast M (2013) Power increases performance in a social evaluation situation as a result of decreased stress responses. *Eur J Soc Psychol*, 43(3): 201-211.
6. Cuddy AJ, Wilmuth CA, Yap AJ, & Carney DR (2015) Preparatory power posing affects nonverbal presence and job interview performance. *J Appl Psychol*, 100(4): 1286.
7. Cuddy AJ, Glick P, & Beninger A (2011) The dynamics of warmth and competence judgments, and their outcomes in organizations. *Res Organ Behavior*, 31: 73-98.
8. Cheng JT, Tracy JL, Foulsham T, Kingstone A, & Henrich J (2013) Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *J Pers Soc Psychol*, 104(1): 103.
9. Mehta PH & Josephs RA (2010) Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Horm Behav*, 58(5): 898-906.
10. Hardy CL & Van Vugt M (2006) Nice guys finish first: The competitive altruism hypothesis. *Pers Soc Psychol Bull*, 32(10): 1402-1413
11. Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull*, 130(3), 355.
12. Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003) Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology*, 28(7): 916-931.
13. Bates D, Maechler, M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw*, 67(1): 1-48.
14. Marceau K, Ruttle PL, Shirtcliff EA, Hastings PD, Klimes-Dougan B, & Zahn-Waxler C (2015) Within-person coupling of changes in cortisol, testosterone, and DHEA across the day in adolescents. *Dev Psychobiol*, 57(6): 654-669.
15. Laurent HK & Powers SI (2006). Social-cognitive predictors of hypothalamic-pituitary-adrenal reactivity to interpersonal conflict in emerging adult couples. *J Soc Pers Relat*, 23(5): 703-720.

16. Stephens MAC, Mahon PB, McCaul ME, & Wand GS (2016) Hypothalamic–pituitary–adrenal axis response to acute psychosocial stress: Effects of biological sex and circulating sex hormones. *Psychoneuroendocrinology*, 66: 47-55.

17. Adler NE, Epel ES, Castellazzo G, Ickovics JR (2000) Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology, 19*, 586-592.

18. Canty A, Ripley B (2016). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-18.

19. Davison, A. C. & Hinkley, D. V. (1997) Bootstrap Methods and Their Applications. *Cambridge University Press, Cambridge*.

20. Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1), 77-89.

21. Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, 8(1), 23.

22. Wirth MM, Schultheiss OC (2006) Effects of affiliation arousal (hope of closeness) and affiliation stress (fear of rejection) on progesterone and cortisol. *Horm Behav*, *50*(5): 786-795.