RESEARCH ARTICLE

REVISED

Stress, rejection, and hormones: Cortisol and progesterone reactivity to laboratory speech and rejection tasks in women and men [v2; ref status: indexed, http://f1000r.es/4l7]

Allison E. Gaffey, Michelle M. Wirth

Department of Psychology, University of Notre Dame, Notre Dame, 46656, USA

v2 **First published:** 02 Sep 2014, **3**:208 (doi: 10.12688/f1000research.5142.1)
Latest published: 30 Oct 2014, **3**:208 (doi: 10.12688/f1000research.5142.2)

Abstract

Stress and social rejection have important impacts on health. Among the mechanisms implicated are hormonal systems such as the hypothalamic-pituitary-adrenal (HPA) axis, which produces cortisol in humans. Current research employs speech stressors and social rejection stressors to understand hormonal responses in a laboratory setting. However, it is not clear whether social rejection stressors elicit hormonal reactivity. In addition to cortisol, progesterone has been highlighted as a potential stress- and affiliation-related hormone in humans. In the present study, 131 participants (70 men and 61 women) were randomly assigned to be exposed to one of four conditions: standardized speech stressor; speech control; social rejection task; or a control (inclusion) version of the social rejection task. Saliva samples were collected throughout the study to measure cortisol and progesterone. As hypothesized, we found the expected increase in cortisol in the speech stressor, and we also found that the social rejection task did not increase cortisol, underscoring the divergence between unpleasant experiences and HPA axis activity. However, we did not find evidence for progesterone increase either during the speech- or social rejection tasks. Compared with past studies on progesterone and stress in humans, the present findings present a mixed picture. Future work is needed to delineate the contexts and types of manipulations which lead to progesterone increases in humans.

Open Peer Review

Referee Status: ✓ ✓

	Invited Referees 1	2
REVISED version 2 published 30 Oct 2014	✓ report	✓ report
version 1 published 02 Sep 2014	? report	? report

1 **Richard Slatcher**, Wayne State University USA

2 **Jens Pruessner**, McGill University Canada

Discuss this article

Comments (0)

Corresponding author: Michelle M. Wirth (mwirth@nd.edu)

How to cite this article: Gaffey AE and Wirth MM. **Stress, rejection, and hormones: Cortisol and progesterone reactivity to laboratory speech and rejection tasks in women and men [v2; ref status: indexed, http://f1000r.es/4l7]** *F1000Research* 2014, **3**:208 (doi: 10.12688/f1000research.5142.2)

Grant information: This research was funded in part by discretionary funds to Michelle Wirth from the University of Notre Dame. An NSF Graduate Student Fellowship supported Allison Gaffey during data analysis and manuscript preparation.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: No competing interests were disclosed.

First published: 02 Sep 2014, **3**:208 (doi: 10.12688/f1000research.5142.1)
First indexed: 05 Jan 2015, **3**:208 (doi: 10.12688/f1000research.5142.2)

REVISED **Amendments from Version 1**

1. Changes to Introduction in order to highlight:
    1. Not just chronically high levels of cortisol, but other kinds of dysregulation in the HPA axis are associated with poorer health;
    2. Loneliness is not identical to rejection, although rejection can engender feelings of loneliness;
    3. The directions of associations between progesterone (or allopregnanolone) and health and behavioral outcomes.
2. Added detail to Methods:
    1. That judges in the TSST stress condition were always one male, one female;
    2. Clarifying that the other 'players' in Cyberball were always the same sex as the participant;
    3. Describing the reason for freeze-thaw cycles as part of saliva processing.
3. Added post-hoc power analysis findings to Results; included implications for the power analysis for our null findings in the Discussion; removed mention of power as a potential limitation.
4. Added to the Discussion: a recommendation that future research examining progesterone levels in women should more carefully control menstrual status.

See referee reports

Introduction

There is a growing interest in human behavioral endocrinology. Encouraged by the availability of non-invasive salivary hormone measurements, researchers in clinical, social, and personality psychology, among other fields, are increasingly incorporating hormonal measurements into their research in order to discover the impact of stress and other kinds of social or emotional stimuli on hormonal systems in human beings.

Among many hormone-relevant psychological constructs, affiliation and bonding, and the converse, isolation and rejection, have received particular attention. Loneliness and lack of social support are known to have grave psychological and health impacts over time (see e.g. Hawkley & Cacioppo, 2010 for a review). Dysregulation in the hypothalamic-pituitary-adrenal (HPA) axis, resulting in e.g. chronically high levels or dyregulated diurnal patterns of glucocorticoids, has been proposed as one possible mechanism mediating the connection between isolation and poor health (Hawkley & Cacioppo, 2010). This idea is supported by evidence that loneliness correlates with higher levels of cortisol, the primary glucocorticoid in humans (Hawkley & Cacioppo, 2010); the fact that social isolation is a potent stressor and elicitor of glucocorticoid release in other social animals, such as rats and sheep (e.g., Hermes *et al.*, 2009; Rivalland *et al.*, 2007); and that dysregulated or chronically high glucocorticoid levels are linked to a number of health consequences (Sapolsky, 2002; Tsigos & Chrousos, 2002). However, the relationships between social isolation, HPA axis activation, and health are complex and still not understood completely – including how acute social rejection, one source of isolation or loneliness, affects physiology. It is necessary to study these relationships on both a macro-level in real-world, longitudinal data (e.g., chronic loneliness/isolation) and also at a micro-level in controlled laboratory settings (e.g., acute social rejection) in order to precisely define the mechanisms involved.

Researchers have used laboratory rejection tasks such as Cyberball — a ball-playing game in which other players exclude the participant — in order to test both the psychological and hormonal effects of social rejection (Maner *et al.*, 2010; Stroud *et al.*, 2002; Williams *et al.*, 2000; Zwolinski, 2012). However, studies have failed to find consistent hormone responses to rejection (Zwolinski, 2012). There has also been evidence of sex differences in hormonal responses to rejection (Stroud *et al.*, 2002), but these effects were not replicated in a separate study (Linnen *et al.*, 2012). It is important to determine whether rejection in a laboratory setting can elicit an HPA axis response, and if so, in which sex or sexes.

Psychological factors known to influence the HPA axis include novelty, unpredictability, and a lack of control (Mason, 1975). A more recent meta-analysis identified social-evaluative threat as key in predicting HPA axis responsivity to laboratory stress tasks (Dickerson & Kemeny, 2004). Any or all of these factors might be present to some degree in a rejection situation, so a cortisol response to rejection in the laboratory could be expected. On the other hand, the main function of glucocorticoids is to mobilize energy, e.g. for fight-or-flight activities (Nelson, 2005; Sapolsky, 2002; Wirth & Gaffey, 2013). Therefore, glucocorticoids do not show a one-to-one relationship with negative affect, but instead are elevated in situations requiring energy, whether associated with negative affect or not; some examples include sickness, exercise, and giving a speech (Wirth *et al.*, 2011; Wirth & Gaffey, 2013). Whereas commonly-used speech stressors require literally thinking on one's feet and making a vigorous (and ultimately futile) attempt to impress the judges, social rejection in laboratory tasks like Cyberball may or may not demand any expenditure of energy – in fact, it may be a situation in which no obvious actions can be taken. Therefore, it is unclear whether laboratory social rejection is a context in which the brain and body would activate a system designed to replenish energy. The first goal of the present study, then, is to examine the effect of a popular rejection manipulation, Cyberball (Williams *et al.*, 2000), on cortisol levels in men and women, alongside the effect of a well-studied, standardized laboratory stressor, the Trier Social Stress Test (TSST; Kirschbaum *et al.*, 1993).

In addition to cortisol, there is a growing body of literature linking progesterone levels/responses to both stress and to affiliation and rejection (Brown *et al.*, 2009; Childs *et al.*, 2010; Gettler *et al.*, 2013; Maner *et al.*, 2010; Schultheiss *et al.*, 2003; Schultheiss *et al.*, 2004; Wirth & Schultheiss, 2006; Wirth *et al.*, 2007; Wirth, 2011). Progesterone is not only a gonadal hormone, but is also produced in the adrenal glands, and progesterone levels increase in response to pharmacological stimulation of the HPA axis (Genazzani *et al.*, 1998). Progesterone and hormones synthesized from it (e.g., allopregnanolone) increase during stress in laboratory animals (Barbaccia *et al.*, 2001; Paul & Purdy, 1992; Purdy *et al.*, 1991), but it is as of yet unclear whether progesterone is part of the typical human stress response (Wirth, 2011). There is evidence that progesterone does increase alongside cortisol during venipuncture stress (Wirth, 2011), and also evidence that progesterone responds to the TSST stressor, at least in men, and in women in some menstrual cycle phases (Childs *et al.*, 2010). Progesterone responses to laboratory stressors need to be studied systematically in both sexes, in part simply to understand stress physiology, but also because of important implications for understanding psychological disorders

(e.g., lower allopregnanolone levels seen in depression; see Wirth, 2011 for a review). Furthermore, progesterone might be particularly associated with affiliation and rejection/isolation, as detailed below.

Although cortisol and progesterone levels seem to rise and fall in tandem in humans (Wirth *et al.*, 2007), a growing body of literature supports associations with affiliation that are unique to progesterone. First, implicit affiliation motivation – a personality construct measuring drive for friendly, warm contact with others - was increased in women taking oral contraceptives containing progestins, as well as in cycling women in the luteal phase, a time in the cycle of high progesterone levels (Schultheiss *et al.*, 2003). Second, a rejection-themed film excerpt designed to produce affiliation-related stress caused increases in progesterone as well as cortisol; in addition, baseline (pre-film) affiliation motivation predicted stress-related increases in progesterone (but not cortisol), without regard to participant sex (Schultheiss *et al.*, 2004; Wirth & Schultheiss, 2006). Third, women who took part in a closeness-generating task in pairs had progesterone increases in response to the task, compared to a control condition (Brown *et al.*, 2009). Fourth, personality traits such as social anxiety and rejection sensitivity moderated progesterone responses to a laboratory rejection task (Maner *et al.*, 2010). Finally, recent, preliminary research links progesterone to the beneficial effects of helping behavior on cardiovascular recovery from stress (Brown & Brown, 2011; Smith, 2011) and to positive mood during fathers' interactions with their toddlers (Gettler *et al.*, 2013).

Given this evidence, along with evidence that progesterone may respond to typical laboratory stressors (Childs *et al.*, 2010; Wirth, 2011), is not yet clear whether progesterone is a "generic" stress hormone, i.e. responding to all stressors along with cortisol, or whether it is tied specifically to affiliation stress/rejection. Notably, in some of the studies cited above, progesterone and not cortisol showed (positive) associations with affiliation (e.g. Wirth & Schultheiss, 2006). Thus, this evidence calls for further research elucidating progesterone's role in stress, affiliation, and rejection. While there is at least one study of progesterone in the context of laboratory rejection tasks (Maner *et al.*, 2010), moderating variables were the focus of that study; more work is needed to determine whether progesterone typically increases during rejection in human beings. Thus, the second goal of the current research is to test whether progesterone increases in response to either the rejection manipulation Cyberball, and/or a standard speech stressor (the TSST).

In both goals of the present research, it is important to determine if there are sex differences. Men typically have larger cortisol responses to laboratory stressors than women do, despite women having equivalent, or even greater, self-reported mood responses (Kudielka & Kirschbaum, 2005). On the other hand, women are thought to be more sensitive to rejection than men (Stroud *et al.*, 2002). In addition to cortisol, progesterone responsivity to both rejection and a speech stressor may have important sex differences (e.g., Childs *et al.*, 2010). For these reasons, we collected data in both women and men exposed to Cyberball or the TSST.

Our hypotheses were four-fold. We expected to (1) replicate substantial prior research (Dickerson & Kemeny, 2004; Kirschbaum *et al.*, 1993; Kudielka & Kirschbaum, 2005) in that the TSST would cause increases in cortisol, particularly in men. We further hypothesized (2) that the TSST would have a greater effect on cortisol than would Cyberball, as the latter is not associated with clear needs for energy mobilization. As for progesterone, we hypothesized that (3) it would increase alongside cortisol in the TSST, as seen in men in at least one previous study (Childs *et al.*, 2010). Given evidence for particular associations with rejection, we also hypothesized that (4) progesterone levels would be affected by Cyberball. We were agnostic as to whether this effect would be present in both sexes, given the paucity of published data on this topic.

Methods

Participants

Undergraduate students ($N = 142$: 71 men: $M_{age} = 19.51$, $SD_{age} = 1.39$; 71 women: $M_{age} = 19.81$, $SD_{age} = 2.43$) were recruited through the University of Notre Dame Psychology Department study pool and through flyers advertising a paid research study open to nonsmoking individuals 18 and 35 years of age. Exclusion criteria included currently nursing or pregnant, and hormonal conditions such as thyroid disorders. In addition, 9 women were taking oral contraceptives and were excluded from analyses. Participants received study pool credit or a cash payment of U.S. $10/hour. The procedures were approved by the University of Notre Dame Institutional Review Board (Protocol #12-09-486), and all participants provided informed consent prior to participation. One man and one woman were excluded from all analyses due to minor changes in the protocol after their participation, leaving a final sample size of 131.

Procedure

Data were collected between October 2010 and July 2011. Participants were asked to refrain from eating, drinking caffeine, brushing their teeth and vigorous exercise for 2 hours prior to the study. Participants completed one session, lasting 150 minutes, between 16:00 and 19:00 to minimize circadian fluctuations in cortisol and progesterone (Dickerson & Kemeny, 2004; Groschl *et al.*, 2003; Hansen *et al.*, 2008; Nelson, 2005). Participants were randomly assigned to one of four conditions: 1) The "stress" condition of the Trier Social Stress Task, including an evaluated speech and difficult serial subtraction (TSST Stress; $N = 36$; Kirschbaum *et al.*, 1993), 2) A "control" version of the TSST during which participants wrote an essay about their dream job and performed a simple addition task alone (without judges; TSST Control; $N = 26$), 3) The "inclusion" condition of Cyberball (Cyberball Control; $N = 32$) or 4) the "rejection" condition of Cyberball (Cyberball Rejection; $N = 37$) (Williams *et al.*, 2000). To match the timing required for the TSST (15 minutes), prior to playing Cyberball, all participants wrote an essay about their dream job for 10 minutes; participants were informed that the essay's content would not be judged or evaluated. The four tasks are further detailed below.

Upon arrival to the laboratory, after obtaining written and verbal consent, participants provided a 5 mL saliva sample (~10 min. after arrival; see saliva collection methodology below) and completed initial questionnaires (~20 min. after arrival). Questionnaires assessed demographic information, affect, and factors that influence hormone levels such as sleep, exercise, and menstrual stage (see Supplementary file). A professional online survey distribution tool, the Qualtrics Survey Research Suite (Qualtrics, Provo, Utah),

was used to capture all self-report data. After completing these initial questionnaires, participants provided a second saliva sample (~30 min.).

Participants were then given directions associated with their randomly assigned task (i.e. Cyberball or TSST) and condition (i.e., Stress/Rejection or Control) before providing a third saliva sample (~50 min.). All participants then engaged in one of the four task-condition combinations. After the Cyberball task, all Cyberball participants completed additional assessments of inclusionary status and ostracism used in previous research (Zadro *et al.*, 2004). Example questions included evaluating the degree to which they "Felt like an outsider during the Cyberball game" and "To what extent did the other participants include you during the game?"

Following the TSST task or Cyberball ostracism questionnaires, participants completed a fourth saliva sample (~70 min.). Participants provided their fifth saliva sample (~105 min.) and sixth and final saliva sample (~150 min.) interspersed among affect questionnaires and non-emotionally-arousing tasks used to test separate hypotheses. Finally, participants completed an open-ended question of any comments or notes about the study, as a suspicion check for Cyberball. The timeline of events in each study session is shown in Figure 1.

Tasks

Trier Social Stress Test (TSST). In the TSST (Kirschbaum *et al.*, 1993), participants have 5 minutes to prepare a speech on a topic they are not well prepared for; in this study they were instructed to try to convince judges who were "experts in judging non-verbal behavior" that they were the best candidate for their dream job. Participants were instructed to only use true information about themselves in their speech. Just before giving their speech, participants' notes were unexpectedly removed. Participants then gave their speech for 5 minutes in front of two judges, always one male and one female, trained to display flat affect (i.e. no smiling or nodding) and give prompts if the participant still had time remaining. Participants also were told they were being videotaped and were able to view themselves on closed-circuit computer monitor. Following the speech, participants completed a 5-minute difficult serial subtraction task out loud for the judges (e.g., count down from 1037 by 13's). The judges required participants to start the task over whenever they made a subtraction mistake. Participants were fully debriefed at the end of the study that they had not, in fact, been videotaped, and that the judges were trained to display flat affect and otherwise increase the stress of the situation, rather than being experts in non-verbal behavior.

Many different control conditions have been used for the TSST (see e.g. Het *et al.*, 2009; Kirschbaum *et al.*, 1993). In the present study, TSST Controls were asked to write an essay about their dream job. Experimenters informed participants in the TSST Control condition that the essay's content would not be judged or evaluated. Additionally, TSST Control participants performed an easy counting task out loud (e.g., count down from 300 by 1's) while alone in the TSST room. Thus, participants in this condition performed the same tasks as in the TSST Stress condition, but without pressure and without being watched or judged.

Cyberball. Cyberball is a computer "ball-toss" game during which participants are either included or ostracized by the other players in order to elicit feelings of social rejection (Williams *et al.*, 2000). Participants in the Cyberball task were randomly assigned to either an inclusion (Control) condition, in which they were passed the ball equally often as the other players, or an exclusion/rejection condition, in which they were passed the ball equally often initially and then excluded from play for the rest of the game. Participants' photographs were taken at the beginning of the session to accompany their character in the Cyberball game. Participants were told that the other two same-sex players (whose behavior was actually computer-generated) were located at another laboratory on campus. Before the game, experimenters made a fake phone call to the fictional lab; this call was intended to be overheard by participants to give the impression that the experimenters were synchronizing Cyberball players' log-ins in the two labs. Names and photographs (students from another university; always both of the same sex as the participant) also accompanied computer players. As a supposed precaution, participants were asked to inform the experimenter at the beginning of the game if they knew the other participants. None of the participants indicated in their final study comments that they did not believe the Cyberball cover story. Participants were fully debriefed at the end of the study that the other players' actions were generated by the computer. Participants played Cyberball for 5 minutes with two other players. The game was set for 100 throws.

Salivary hormone measures

We assessed cortisol and progesterone levels in the six saliva samples provided by each participant. Participants used passive drool

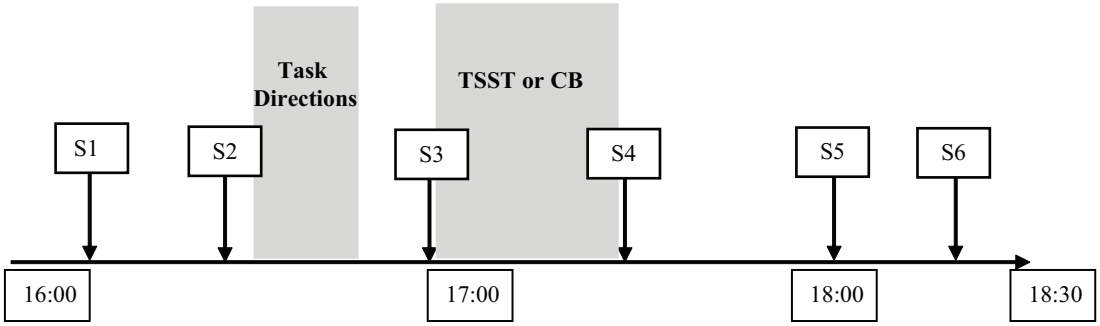

Figure 1. Study timeline. S1, S2, etc. represent saliva samples; approximate times are shown for the study session on a 24-hour clock.

into a straw (i.e., no gum, cotton, or other saliva flow stimulants) to deposit saliva into a test tube (typically, Ultra-High Performance 15 ml centrifuge tubes, VWR, Radnor, PA), and were allowed to drink sips of water following each sample. Tubes were capped and frozen at -18°C after each data collection session. After sample collection, saliva samples underwent three freeze-thaw cycles (i.e. samples were thawed until liquid and re-frozen until solid, twice) in order to break up mucopolysaccharides and reduce viscosity to aid in accurate pipetting, followed by centrifugation (10 min at 3000 rpm). Cortisol and progesterone levels were determined by solid-phase ^{125}I radioimmunoassays (Coat-A-Count, Siemens Healthcare Diagnostics, Duluth, GA), using the protocol described by Wirth & Schultheiss, (2006). Range of standards used was 0.5 to 50 ng/ml for cortisol and 5 to 400 pg/ml (i.e., 0.005 to 0.4 ng/ml) for progesterone. A total of 8 assays for each hormone were performed in order to assay all 852 samples. Mean intra-assay coefficients of variation (CV) across all 852 samples were 7.1% for cortisol and 19.9% for progesterone. (Since progesterone is present at much lower concentrations than cortisol, CVs are typically much higher than for cortisol; see e.g. [Wirth & Schultheiss, 2006]. Average CVs for progesterone in this range have been reported in the literature previously and have been associated with theoretically-supported positive findings [Brown *et al.*, 2009]). Inter-assay CVs for Stress and Control combined pools of saliva averaged 5.3% and 1.9% for cortisol, and 8.4% and 10.1% for progesterone. Averaged across the 8 assays, the lower limit of detection (B_0 – 3 x SD method) was 0.1 ng/ml for cortisol assays and 3.9 pg/ml for progesterone assays. Average recovery values for external controls (Lyphocheks) were 90.2 and 90.8% for low and high concentration in progesterone assays, and 119.5 and 119.1% for low and high concentration cortisol controls.

Data analysis

Data were analyzed using SYSTAT 13 and SPSS 21. Where raw hormone data are presented, salivary cortisol concentrations are reported as ng/ml and progesterone concentrations as pg/ml. To examine the overall magnitude of hormonal response to the tasks, we calculated the area under the curve with respect to increase (AUC_i; Pruessner *et al.*, 2003) from cortisol and progesterone Sample 3 (baseline/pre-task) to Sample 6 (post-task, at the end of the study). Sample 3 was chosen as the baseline as stress hormones are well-known to be elevated at the beginning of study sessions, owing to the novelty of the test environment, among other factors (see e.g. cortisol data in Abercrombie *et al.*, 2006; further explanation in Wirth *et al.*, 2011). Notably, AUC_i calculations improve on difference scores because they utilize information for all measurements from Sample 3 to Sample 6. Previous studies have shown that cortisol tends to be elevated for up to 90 minutes after the TSST (e.g., Kirschbaum *et al.*, 1995), so Sample 6 is timed appropriately to capture the end of most hormonal responses to the task. Therefore, the chosen number of samples and the timeframe used to calculate AUC_i were selected to capture the complete cycle of hormonal change in response to the stressors/tasks.

To test our hypotheses about effects of the manipulations, as well as to test for sex differences, we first conducted an ANOVA on AUC_i for the entire sample, with Group (TSST stress, TSST control, Cyberball rejection, or Cyberball control) and Sex as the independent variables. Second, to further explore how the effects emerge for each sex, we split the sample by sex and conducted ANOVAs for each sex on AUC_i by group. Post-hoc Tukey tests were then used to follow up on all ANOVAs.

Menstrual phase could be expected to impact hormone levels, particularly progesterone. Fortunately, by using AUC, initial differences in progesterone due to variations in menstrual phase are controlled for, since AUC_i reflects the total amount of increase in the hormone from baseline — in other words, baseline differences are factored out. Furthermore, there was no correlation between progesterone AUC_i and self-reported number of days since the start of the last menstrual period, i.e. the point that each woman was in her cycle (r^2 = -0.077, p = 0.57). Neither was this relationship significant for cortisol AUC_i (r^2 = -0.073, p = 0.59). Also, self-reported days since period, entered as a covariate, did not moderate the effect of Group on either cortisol AUC_i or progesterone AUC_i in women. Therefore, for the purposes of the present research, we conducted analyses collapsing over menstrual phase. To more directly address the question of how menstrual phase impacts hormonal responses to tasks like the TSST and Cyberball, research would be needed selecting women in particular cycle phases; this was beyond the scope of the present report.

Results

Power analysis

We performed a post-hoc power analysis using G*Power 3.1 to determine whether we had achieved adequate power to detect the small effect size we obtained (see below) in an overall F test by Group and Sex on cortisol response (AUC_i). Using our obtained partial eta squared of 0.10 (i.e., a small effect size), for a 2-way ANOVA with 8 total groups and a sample size of 131, we had power of 0.90 to detect an effect of this size. Therefore, we feel confident that the study was adequately powered.

Cyberball manipulation check

Participants who completed Cyberball, in both the inclusion and exclusion conditions, completed a questionnaire afterwards rating a number of statements regarding their inclusion and feelings during the game (Williams *et al.*, 2000). T-tests were used to compare participants' ratings on these items in the inclusion (control) vs. exclusion (stress) condition. As expected, participants in the exclusion condition rated that a smaller percent of the throws were made to them, and that the other game-players included them less, as well as excluded them more. They were less likely to endorse that they made a connection or bonded with one or more of the other game-players, and they rated themselves as feeling more like an outsider, more non-existent, and less in control. They rated themselves as feeling less able to throw the ball as often as they wanted, and less that their performance had any effect on the direction of the game. They also were significantly more likely to endorse that the other game-players failed to perceive them as worthy and likeable people (all $p < 0.05$). Excluded participants also endorsed at marginally greater rates the statement "I felt somewhat inadequate during the Cyberball game" (p = 0.055). There were no significant differences between the exclusion and inclusion groups on statements regarding feeling frustrated, angry, good about oneself, enjoyment of the game, or "felt as though my existence was meaningless" (even though excluded participants did rate "I felt non-existent during the game" significantly higher than included participants). Therefore, participants were clearly aware of the exclusion and had negative feelings about it. As

mentioned above, when given an opportunity to give comments or observations about the study, no participants expressed suspicion that the other game-players were not real people.

Cortisol

An ANOVA with factors Group and Sex yielded a significant main effect of Group on cortisol AUC_i, $F(3,130) = 4.54$, $p = 0.005$, *partial* $\eta^2 = 0.100$. Neither the main effect of Sex nor the interaction was significant (main effect of sex: $p = 0.188$; interaction: $p = 0.698$). As expected, cortisol AUC_i was highest in the TSST Stress group: *M* (*SD*) = 9.26 (32.67), compared with -2.59 (26.72) for TSST Control; and -16.41 (37.44) and -4.25 (16.52) for Cyberball Rejection and Control, respectively. Post-hoc Tukey HSD tests by Group revealed a significant pair-wise comparison only between TSST Stress and Cyberball Rejection groups ($t(71) = -3.12$, $p = 0.003$; *95% CI:* -42.094 to -9.260, *Cohen's d* = -0.731). However, as seen by the mean AUCs, the TSST Stress group was the only group with a positive AUC_i, reflecting an overall increase in cortisol over the session.

In exploratory, separate ANOVAs conducted in women and men, Group significantly impacted cortisol AUC_i in men ($F(3,69) = 3.86$, $p = 0.013$, *partial* $\eta^2 = 0.149$). Post-hoc Tukey tests in men again revealed a significant pair-wise comparison between TSST Stress and Cyberball Rejection ($t(35) = -2.70$, $p = 0.011$; *95% CI:* -43.53 to -6.18, *Cohen's d* = -0.883) as well as a significant comparison between TSST Stress and TSST Control ($t(32) = 2.07$, $p = 0.046$; *95% CI:* 0.325 to 37.656). In women, though again the highest and only positive cortisol AUC_i was in the TSST Stress group (*M* (*SD*) AUC_i = 3.23 (31.17), vs. -0.823 (40.23) in TSST Control; -23.65 (48.97) in Cyberball Rejection; -6.89 (15.49) in Cyberball Control), the ANOVA in women failed to reach significance ($F(3,60) = 1.82$, $p = 0.154$, *partial* $\eta^2 = 0.087$). See Figure 2.

In sum, TSST and Cyberball do not have the same effects on cortisol levels. The TSST Stress condition was the only condition which caused an increase in cortisol. Sex did not moderate this finding; however, when the sample was split by sex in an exploratory analysis, only in men did the effect remain significant.

Progesterone

An ANOVA conducted on progesterone AUC_i with factors Group and Sex yielded no significant main effects or interactions, all $p > 0.4$. Neither were there any effects of Group on progesterone AUC_i when examined separately in men or in women. Interestingly, in men, the Cyberball Rejection condition elicited the highest average progesterone AUC_i out of the four groups; mean AUC_i in men in Cyberball Rejection was 60.04 (229.55), vs. -45.20 (293.55) in Cyberball Control, and -69.37 (425.20) and 14.02 (353.24) in TSST Stress and Control, respectively. However, pairwise post-hoc comparisons failed to reach significance. Thus, there were no effects of either Cyberball or TSST on progesterone in either sex; see Figure 3.

Cortisol and progesterone data collected in participants exposed to speech and rejection tasks

1 Data File

http://dx.doi.org/10.6084/m9.figshare.1150167

Discussion

This study evaluated the effects of two different stress tasks, and their respective controls, on cortisol and progesterone. We found support for our first and second hypotheses, in that the TSST elicited a significantly greater cortisol response than all other tasks. Cyberball exclusion/social rejection was not associated with cortisol reactivity; we can be fairly confident in this null finding given results of our power analysis. Our cortisol findings are in line with the physiological functions of glucocorticoids, which include mobilizing energy (Nelson, 2005; Sapolsky, 2002; Wirth & Gaffey, 2013). Cyberball exclusion is certainly unpleasant for participants (Williams *et al.*, 2000; Zadro *et al.*, 2004), but it is not a situation that demands or even allows very much active thought, planning, or physical activity. This is in contrast to the TSST, in which participants are continually actively modifying their speech in response to the feedback (or lack thereof) from the judges. The performance aspect of the TSST possibly requires more energy consumption by both the brain and body, and therefore a higher glucocorticoid response compared with Cyberball, which involves simply sitting at a computer pressing keys to determine the direction of the next ball toss.

These findings also underscore the fact that not every situation involving social rejection and associated negative feelings engenders a cortisol response, as well as the lack of a one-to-one relationship between negative feelings/mood/affect and cortisol. There are many examples of conditions in which cortisol is elevated without necessarily any changes to mood or affect, including exercise and illness. There are also examples of laboratory stimuli which sharply increase negative affect without affecting cortisol levels, such as viewing unpleasant pictures (Wirth *et al.*, 2011; Wirth & Gaffey, 2013). Furthermore, meta-analyses across laboratory stressors show small or zero correlations between cortisol and subjective emotional responses (Campbell & Ehlert, 2012; Dickerson & Kemeny, 2004; Page-Gould *et al.*, 2013).

The greater cortisol response to the TSST is also in line with Dickerson & Kemeny's (2004) demonstration that social-evaluative judgment is the key factor in generation of cortisol responses in psychological laboratory tasks. Cyberball might be thought of as including social judgment, but there is very little for the other "players" to judge about the participant. In fact, in Cyberball exclusion, it is completely ambiguous why the other players cease throwing the ball to the participant. In the TSST, on the other hand, the constant monitoring and interruptions of the judges, along with their flat affect, can be taken by a participant to directly relate to their speech and arithmetic performance in real time.

These findings have implications for understanding the health consequences of real-world loneliness and social rejection. It is often speculated that HPA axis activity, specifically higher cortisol levels, might mediate the connection between social rejection and poorer health. However, at least in a laboratory setting, an acute social rejection experience does not cause a cortisol response, suggesting other mechanisms. Alternately, it may be that HPA activity only plays a role in chronic or "real-life" rather than acute, laboratory experiences of social rejection, e.g. loneliness (Adam *et al.*, 2006; Hawkley & Cacioppo, 2010). Cyberball may not be the ideal task to study social rejection in the laboratory in relation to detrimental effects on health.

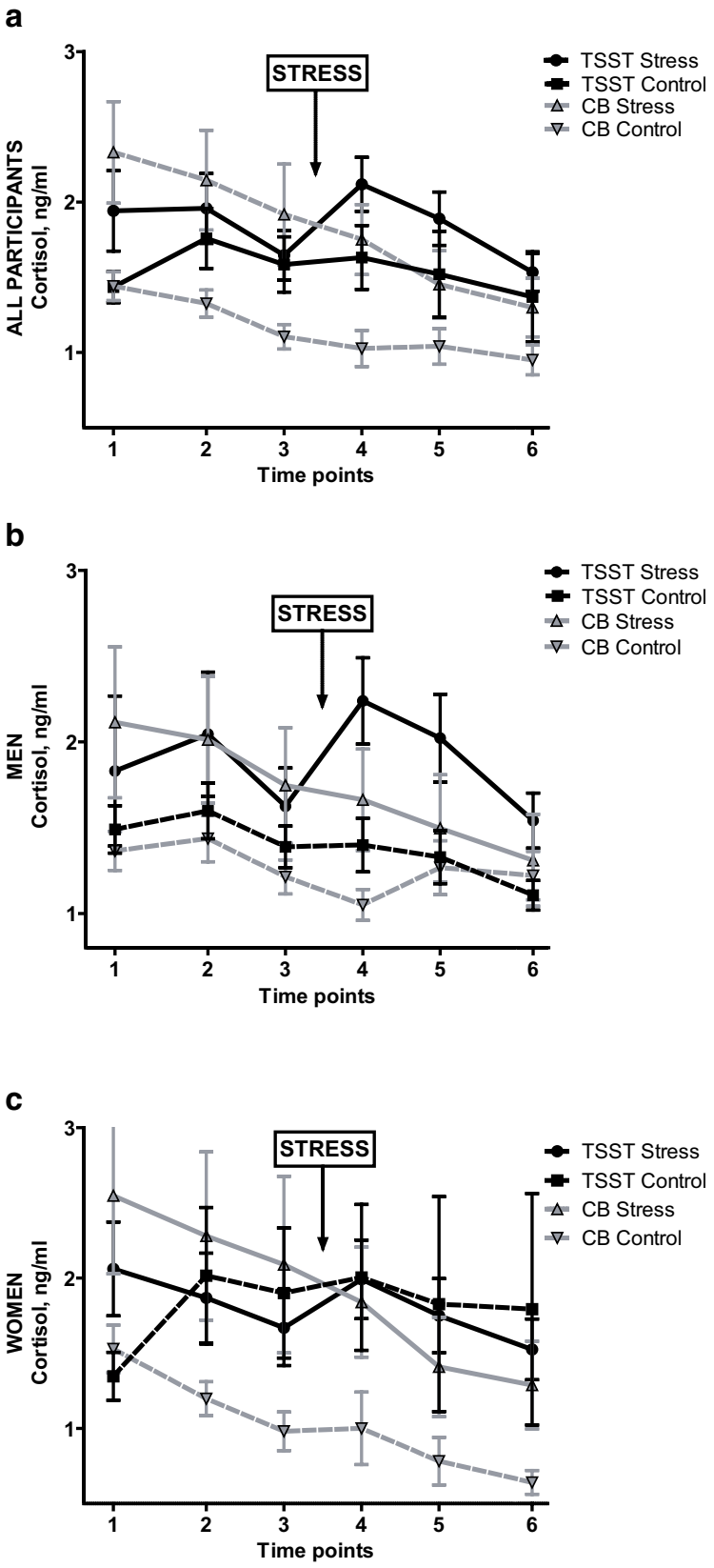

Figure 2. Salivary cortisol by condition. Salivary cortisol for: entire sample, (**a**); men only, (**b**); and women only, (**c**). TSST = Trier Social Stress Test. CB = Cyberball. Ng/ml = nanograms per milliliter. Error bars indicate standard error of the mean.

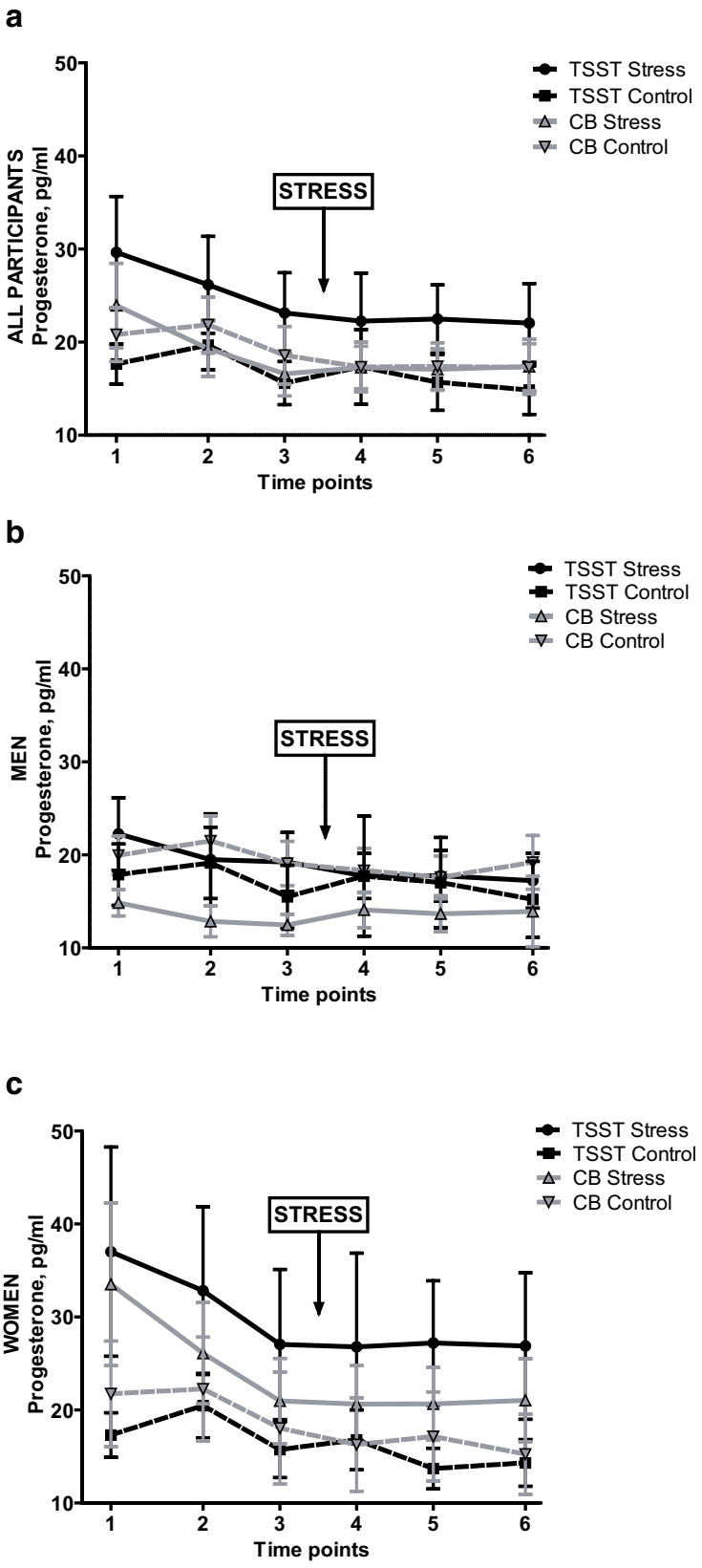

Figure 3. Salivary progesterone by condition. Salivary progesterone for: entire sample, (**a**); men only, (**b**); and women only, (**c**). TSST = Trier Social Stress Test. CB = Cyberball. Pg/ml = picograms per milliliter. Error bars indicate standard error of the mean.

In contrast with our cortisol results, neither the TSST nor Cyberball induced a change in progesterone. This finding is somewhat surprising in light of research demonstrating that progesterone does increase in response to some types of stress (Childs *et al.*, 2010; Wirth, 2011), including social rejection (Maner *et al.*, 2010; Wirth & Schultheiss, 2006). From our power analysis, we can be confident in this null finding to the extent that we can expect a similar (small) effect size in progesterone as we found in cortisol. Regarding our prior reports that progesterone and cortisol levels increase and decrease in tandem in men and in women taking hormonal contraceptives, indicative of progesterone increasing alongside cortisol during stress, it is worth noting that this was not found in cycling women (i.e., women not on hormonal contraceptives, such as in the present study; Wirth *et al.*, 2007). Possibly, progesterone only increases during certain kinds of stressors, such as those including physical pain/distress, such as venipuncture (Wirth, 2011), or only under certain conditions, such as the morning (Childs *et al.*, 2010). Another possibility is that, in social rejection contexts, progesterone responses are driven by a "tend-and-befriend", affiliative response (Wirth, 2011). Though it creates a sense of rejection, the lack of face-to-face contact might cause Cyberball to not generate affiliative motivation to the same extent as other rejection tasks, or even film clips (Wirth & Schultheiss, 2006). Further research is necessary to comprehensively chart under what circumstances and what types of stressors cause increases in progesterone in humans. It is also important to characterize the conditions that provoke increases in downstream hormones like allopregnanolone, since allopregnanolone and related progesterone-derived neurosteroids could be important components of stress regulation (Wirth, 2011).

The limitations of this study should be acknowledged. Logistics of running the study demanded a lack of precise control over which menstrual phase the women participants were in. As mentioned above, however, a self-report measure of menstrual phase did not correlate with AUC_i for either hormone and did not moderate any of the findings. Nonetheless, we recommend that future research assessing progesterone levels in women should more carefully control for menstrual phase/status. A second potential limitation is that, although every effort was made to conceal information about condition/group assignment from the participant until directly before their task, the study was only single-blind, and it is conceivable that the experimenters unconsciously treated stress versus control participants differently prior to the experimental manipulation. Finally, as discussed previously, Cyberball may not be a strong enough manipulation to generalize our findings to any acute social rejection experience.

In conclusion, we found evidence that, unlike a standardized speech task, Cyberball social rejection is not associated with a cortisol response in a sample of college students, despite exclusion-related feelings engendered by this task. This evidence underscores the fact that the HPA axis does not have a one-to-one relationship with social rejection experiences and associated feelings. We also found a lack of evidence for a progesterone response to the cortisol-provoking speech stressor, as well as to the Cyberball rejection task. Taken with past work (Childs *et al.*, 2010; Maner *et al.*, 2010; Schultheiss *et al.*, 2003; Wirth & Schultheiss, 2006; Wirth, 2011), these findings present a mixed picture in terms of evidence for progesterone responsivity to stress (and specifically to social rejection) in humans. Future work is needed to delineate precisely the types of emotional and social manipulations and physical stressors which lead to progesterone increases, as well as downstream neurosteroids. This work is important both from the perspective of basic physiology and psychology research, to understand the hormonal effects of stress and emotion in human beings, and also from a health standpoint, to better understand the mechanisms underlying impacts of stress and social rejection on human health.

Data availability

figshare: Cortisol and progesterone data collected in participants exposed to speech and rejection tasks. doi: 10.6084/m9.figshare.1150167 (Gaffey & Wirth, 2014).

Author contributions

AG and MW together conceived the idea for the study and designed the experiment. AG carried out the research, conducted data analysis, and wrote the initial draft of the manuscript. AG and MW together completed additional writing, edited and finalized the manuscript.

Competing interests

No competing interests were disclosed.

Grant information

This research was funded in part by discretionary funds to Michelle Wirth from the University of Notre Dame. A National Science Foundation (NSF) Graduate Student Fellowship supported Allison Gaffey during data analysis and manuscript preparation.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors would like to thank undergraduate research assistants Christina Buchanan, Kelsey Christoffel, Cristina Kline-Quiroz, Brandy Martinez, Kelly Miller, Kathleen Poplowski, and Mark Tancredi for their assistance with data collection. We thank Brandy Martinez for her assistance with coordinating research assistants, recruitment, data processing, and hormone assays. We also thank Kim Wallen and Joyce Pang for helpful feedback on prior drafts of this manuscript.

Supplementary file

Health Intake survey

Intake questionnaire used to assess demographic information, and factors that influence hormone levels.

Click here to access the data.

References

Abercrombie HC, Speck NS, Monticelli RM: **Endogenous cortisol elevations are related to memory facilitation only in individuals who are emotionally aroused.** *Psychoneuroendocrinology.* 2006; **31**(2): 187–96.
PubMed Abstract | Publisher Full Text

Adam EK, Hawkley LC, Kudielka BM, *et al.*: **Day-to-day dynamics of experience—cortisol associations in a population-based sample of older adults.** *Proc Natl Acad Sci U S A.* 2006; **103**(45): 17058–63.
PubMed Abstract | Publisher Full Text | Free Full Text

Aron A, Aron EN, Smollan D: **Inclusion of other in the self scale and the structure of interpersonal closeness.** *J Pers Soc Psychol.* 1992; **63**(4): 596–612.
Publisher Full Text

Barbaccia ML, Serra M, Purdy RH, *et al.*: **Stress and neuroactive steroids.** *Int Rev Neurobiol.* 2001; **46**: 243–272.
PubMed Abstract | Publisher Full Text

Brown SL, Brown RM: **Activation of the caregiving system: Progesterone as a possible mechanism for linking compassionate behavior to enhanced stress regulation**. Presentation at the *Association for Psychological Science 23rd Annual Meeting*, Washington, D.C. 2011.

Brown SL, Fredrickson BL, Wirth MM, *et al.*: **Social closeness increases salivary progesterone in humans.** *Horm Behav.* 2009; **56**(1): 108–111.
PubMed Abstract | Publisher Full Text | Free Full Text

Campbell J, Ehlert U: **Acute psychosocial stress: does the emotional stress response correspond with physiological responses?** *Psychoneuroendocrinology.* 2012; **37**(8): 1111–1134.
PubMed Abstract | Publisher Full Text

Childs E, Dlugos A, De Wit H: **Cardiovascular, hormonal, and emotional responses to the TSST in relation to sex and menstrual cycle phase.** *Psychophysiology.* 2010; **47**(3): 550–559.
PubMed Abstract | Publisher Full Text

Dickerson SS, Kemeny ME: **Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research.** *Psychol Bull.* 2004; **130**(3): 355–391.
PubMed Abstract | Publisher Full Text

Gaffey A, Wirth M: **Cortisol and progesterone data collected in participants exposed to speech and rejection tasks.** *figshare.* 2014.
Data Source

Genazzani AR, Petraglia F, Bernardi F, *et al.*: **Circulating levels of allopregnanolone in humans: gender, age and endocrine influences.** *J Clin Endocrinol and Metab.* 1998; **83**(6): 2099–2103.
PubMed Abstract

Gettler LT, McDade TW, Agustin SS, *et al.*: **Progesterone and estrogen responsiveness to father-toddler interaction.** *Am J Hum Biol.* 2013; **25**(4): 491–498.
PubMed Abstract | Publisher Full Text

Groschl M, Rauh M, Dorr HG: **Circadian rhythm of salivary cortisol, 17alpha-hydroxyprogesterone, and progesterone in healthy children.** *Clin Chem.* 2003; **49**(10): 1688–1691.
PubMed Abstract | Publisher Full Text

Hansen AM, Garde AH, Persson R: **Sources of biological and methodological variation in salivary cortisol and their impact on measurement among healthy adults: a review.** *Scand J Clin Lab Invest.* 2008; **68**(6): 448–458.
PubMed Abstract | Publisher Full Text

Hawkley LC, Cacioppo JT: **Loneliness matters: a theoretical and empirical review of consequences and mechanisms.** *Annals Behav Med.* 2010; **40**(2): 218–227.
PubMed Abstract | Publisher Full Text | Free Full Text

Hermes GL, Delgado B, Tretiakova M, *et al.*: **Social isolation dysregulates endocrine and behavioral stress while increasing malignant burden of spontaneous mammary tumors.** *Proc Natl Acad Sci U S A.* 2009; **106**(52): 22393–22398.
PubMed Abstract | Publisher Full Text | Free Full Text

Het S, Rohleder N, Schoofs D, *et al.*: **Neuroendocrine and psychometric evaluation of a placebo version of the 'Trier Social Stress Test'.** *Psychoneuroendocrinology.* 2009; **34**(7): 1075–1086.
PubMed Abstract | Publisher Full Text

Kirschbaum C, Klauer T, Filipp SH, *et al.*: **Sex-specific effects of social support on cortisol and subjective responses to acute psychological stress.** *Psychosom Med.* 1995; **57**(1): 23–31.
PubMed Abstract

Kirschbaum C, Pirke KM, Hellhammer DH: **The 'Trier Social Stress Test'—a tool for investigating psychobiological stress responses in a laboratory setting.** *Neuropsychobiology.* 1993; **28**(1–2): 76–81.
PubMed Abstract | Publisher Full Text

Kudielka BM, Kirschbaum C: **Sex differences in HPA axis responses to stress: a review.** *Biol Psychol.* 2005; **69**(1): 113–132.
PubMed Abstract | Publisher Full Text

Linnen AM, Ellenbogen MA, Cardoso C, *et al.*: **Intranasal oxytocin and salivary cortisol concentrations during social rejection in university students.** *Stress.* 2012; **15**(4): 393–402.
PubMed Abstract | Publisher Full Text

Maner JK, Miller SL, Schmidt NB, *et al.*: **The endocrinology of exclusion: rejection elicits motivationally tuned changes in progesterone.** *Psychol Sci.* 2010; **21**(4): 581–588.
PubMed Abstract | Publisher Full Text

Mason JW: **Emotion as reflected in patterns of endocrine integration**. In L. Levi (Ed.), *Emotions - their parameters and measurement.* New York, NY: Raven. 1975; pp. 143–181.

Nelson RJ: Stress. In: **An Introduction to Beh Endocrinology**. 3rd ed,. Sunderland, MA: Sinauer Associates, 2005; pp. 669–720.
Reference Source

Page-Gould E, Khoury J, Fournier M: **The fickle relationship between subjective and hormonal stress**. Presentation at the *2nd Annual Social Neuroendocrinology Preconference of the Society of Personality and Social Psychology*, January 2013, New Orleans, LA. 2013.

Paul SM, Purdy RH: **Neuroactive steroids.** *FASEB J.* 1992; **6**(6): 2311–2322.
PubMed Abstract

Pruessner JC, Kirschbaum C, Meinlschmid G, *et al.*: **Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change.** *Psychoneuroendocrinology.* 2003; **28**(7): 916–931.
PubMed Abstract | Publisher Full Text

Purdy RH, Morrow AL, Moore PH Jr, *et al.*: **Stress-induced elevations of gamma-aminobutyric acid type A receptor-active steroids in the rat brain.** *Proc Natl Acad Sci U.S.A.* 1991; **88**(10): 4553–4557.
PubMed Abstract | Publisher Full Text | Free Full Text

Rivalland ET, Clarke IJ, Turner AI, *et al.*: **Isolation and restraint stress results in differential activation of corticotrophin-releasing hormone and arginine vasopressin neurons in sheep.** *Neuroscience.* 2007; **145**(3): 1048–1058.
PubMed Abstract | Publisher Full Text

Sapolsky RM: **Endocrinology of the stress-response**. In: Becker, J. B., Breedlove, S. M., Crews, D. and McCarthy, M. M. (Eds.), *Behavioral endocrinology 2nd ed.*, pp. 409–450. Cambridge, MA: M.I.T. Press. 2002.
Reference Source

Schultheiss OC, Dargel A, Rohde W: **Implicit motives and gonadal steroid hormones: Effects of menstrual cycle phase, oral contraceptive use and relationship status.** *Horm Behav.* 2003; **43**(2): 293–301.
PubMed Abstract | Publisher Full Text

Schultheiss OC, Wirth MM, Stanton SJ: **Effects of affiliation and power motivation arousal on salivary progesterone and testosterone.** *Horm Behav.* 2004; **46**(5): 592–599.
PubMed Abstract | Publisher Full Text

Smith DM: **Helping others as a protective factor that promotes physiological resilience to stress**. *Association for Psychological Science 23rd Annual Meeting*, Washington, D.C. 2011.

Stroud LR, Salovey P, Epel ES: **Sex differences in stress responses: social rejection versus achievement stress.** *Biol Psychiatry.* 2002; **52**(4): 318–327.
PubMed Abstract | Publisher Full Text

Tsigos C, Chrousos GP: **Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress.** *J Psychosom Res.* 2002; **53**(4): 865–871.
PubMed Abstract | Publisher Full Text

Williams KD, Cheung CK, Choi W: **Cyberostracism: effects of being ignored over the Internet.** *J Pers Soc Psychol.* 2000; **79**(5): 748–762.
PubMed Abstract | Publisher Full Text

Wirth MM: **Beyond the HPA Axis: Progesterone-Derived Neuroactive Steroids in Human Stress and Emotion.** *Frontiers Endocrinol (Lausanne).* 2011; **2**: 19.
PubMed Abstract | Publisher Full Text | Free Full Text

Wirth MM, Gaffey AE: **Hormones and emotion: Stress and beyond**. In: Robinson, M.D., Watkins, E.R. and Harmon-Jones, E. (Eds.), *Handbook of cognition and emotion*, New York, NY: The Guilford Press. 2013; pp. 69–94.
Reference Source

Wirth MM, Meier EA, Fredrickson BL, *et al.*: **Relationship between salivary cortisol and progesterone levels in humans.** *Biol Psychol.* 2007; **74**(1): 104–107.
PubMed Abstract | Publisher Full Text

Wirth MM, Scherer SM, Hoks RM, *et al.*: **The effect of cortisol on emotional responses depends on order of cortisol and placebo administration in a within-subjects design.** *Psychoneuroendocrinology.* 2011; **36**(7): 945–954.
PubMed Abstract | Publisher Full Text | Free Full Text

Wirth MM, Schultheiss OC: **Effects of affiliation arousal (hope of closeness) and affiliation stress (fear of rejection) on progesterone and cortisol.** *Horm Behav.* 2006; **50**(5): 786–795.
PubMed Abstract | Publisher Full Text

Zadro L, Williams KD, Richardson R: **How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence.** *J Exp Soc Psychol.* 2004; **40**(4): 560–567.
Publisher Full Text

Zwolinski J: **Psychological and Neuroendocrine Reactivity to Ostracism.** *Aggress Behav.* 2012; **38**(2): 108–125.
PubMed Abstract | Publisher Full Text

Open Peer Review

Current Referee Status:

Version 2

Referee Report 05 January 2015

doi:10.5256/f1000research.5947.r6607

Jens Pruessner
Departments of Psychiatry, Psychology, Neurology and Neurosurgery, McGill University, Verdun, QC, Canada

I would like to thank the authors for addressing all points raised in response to the original version. The manuscript has been mostly improved as a result, with the added power analysis, comments on the menstrual cycle phase, clarification of the sex of the panel of judges, and clarification for the reason of the repeated freeze-thaw cycles.

A few comments remain, which I'd like to list simply as points of reference for the authors, and readers. I am happy to approve the manuscript as is without reservation, but think it is useful to still communicate these comments as part of the manuscript to allow the reader to take them into consideration.

1. The authors have added a power analysis, but the way the power analysis was conducted and presented is probably not 100% optimal. To use the effect sizes from your own sample for power analysis allows you to indicate how strong the effects were for those tests which came out significant, and how much variability was explained by the factors you manipulated and controlled. That is one aspect of power analysis.

   What might be important to address in addition, however, is the question of what the chances were to observe a significant effect in your sample if it exists in the population, for those tests that were not significant in your study. In the context of the current study, two of the main effects in question did not show significance - a potential progesterone response to stress, and the effect of cyberball on cortisol. Thus, it would be important to indicate what the chances of the current study were to observe significant effects here if they existed in the population.

   I had suggested to use previous studies to estimate the expected effect size in their study and then calculate the power from that for your study, given your alpha-level and your sample size - in the current study, this would translate into finding previous studies that observed a progesterone response to stress, and previous studies that found an effect of the cyberball stressor on cortisol release, estimate the effect size from those studies, and use it to calculate the power to find an effect in the current study. Related to this, it is not sufficient to use one effect size from your own findings for all power analyses (in this case, the group by sex effect on cortisol) - you would want to conduct different power analyses with different effect size estimates depending on the effects under investigation. (See, for example, the article by O'Keefe on that topic: O'Keefe, D. (2007).

2. The authors were not very enthusiastic about the idea of using a within-subject design; this was a side point as their between-subjects design is certainly valid. However, the argument why a

within-subject design might be suboptimal was not entirely convincing to me, either. While habituation to repeated exposure of the same stressor is a significant issue in stress studies, cross-habituation to laboratory stressors in general is less frequently investigated, and - at least to my knowledge - not frequently observed. In fact, the opposite phenomena, sensitization across different stressors, is more prominently investigated and discussed in the stress literature (which a simple search on PubMed using these keywords will reveal). While this would also add to the complexity of interpreting the results, randomization in the order of stressor presentation would help to interpret the results in either case. I certainly do agree with the authors that a full within subject design is impractical for the resulting 16 different orders. But a mixed within and between subject design would probably be feasible, reduce the number of cells substantially, and allow to co-investigate the TSST and the Cyberball task within the same person.

3. The menstrual cycle phase effects might go beyond the effects on changing hormone levels, also influencing what women perceive as stressful depending on the phase they are in (see the recent paper by Duchesne *et al*. (2013) on this topic). Thus, I think it is important to not only control for the baseline levels of hormones affected by the cycle, but ideally also keep the phase constant, or have enough subjects to include the phase as separate factor in the analysis. I thus am grateful that the authors have included a respective comment in the Limitations' section.

4. I thank the authors for clarifying that the increase between time point 3 and time point 4 in the women occurred in the TSST stress group. Regardless, the results depicted in Figure 2 still seem to suggest that the control version of the TSST also led to an - at least descriptive - increase in cortisol levels in the women between samples 3 and 4. I think this suggests that there are some aspects to the control version of the TSST that are perceived as stressful by at least some of the female participants - otherwise, you would expect to observe simply a decline of cortisol levels throughout the task, in line with the circadian rhythm of cortisol. As it is not clear what aspect of the TSST control version leads to the increase in cortisol, it would thus be advisable to not repeat any aspect of this task in any other task, as it otherwise might create a confound. The authors point out that the TSST control task is frequently used - I don't doubt that, but would argue that it doesn't matter. Repeating any aspect of one stress task in another will complicate the interpretation of the results, if the aim is to understand what is stressful about one task vs. another.

5. Thank you for clarifying that the judges' panel in the TSST was always sex/gender mixed. I think this is best practice to avoid any potential effect of the sex/gender of the panel on the results. Along the same lines, the fact that the Cyberball panel was always of the (perceived) same sex/gender raises the question of whether different results could be expected if the panel was of mixed, or opposite sex/gender.

6. Thank you for clarifying why three freeze-thaw cycles were employed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 06 November 2014

doi:10.5256/f1000research.5947.r6606

Richard Slatcher
Department of Psychology, Wayne State University, Detroit, MI, USA

The authors did an excellent job addressing my comments and concerns.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Version 1

Referee Report 06 October 2014

doi:10.5256/f1000research.5482.r6015

Jens Pruessner
Departments of Psychiatry, Psychology, Neurology and Neurosurgery, McGill University, Verdun, QC, Canada

The paper by Gaffey and Wirth pursued two main goals. One, to identify whether 'Cyberball', a frequently used task in the Social Neurosciences, elicits a significant cortisol response (and how that compares to the well described cortisol stress response of the Trier Social Stress Test, TSST). Two, to determine whether these stress tasks also elicit a significant change in circulating levels of progesterone, as previous research provided some evidence that progesterone might respond to a psychosocial stress task as well.

The authors report that there was no increase in cortisol to the 'Cyberball' task (although there was the expected increase in response to the TSST), and they further did not observe any significant change in progesterone levels in response to either task.

There are many positive things to be said about the study. The experimental design is innovative, and the major research questions (both relating to the stressfulness of Cyberball and the responsivity of progesterone to stress) are original. The combination of the two goals into one study is actually an added plus, as this allowed the cross-validation of a possible progesterone response to a stressful task using two different experimental manipulations.

The study is also nicely powered (n=131), although not completely balanced. The obvious main drawback of the study is the lack of significant results; unfortunately, the authors could not find any evidence for a progesterone response to stress, and no evidence for a cortisol response to stress. As with any negative finding, it would be important to actually estimate the expected effect size from previous studies and use that to compute the power for the current study - in other words, what were the chances of finding a significant effect given the sample size present in this experiment? That information would allow the reader some judgment as to whether no significant changes are to be expected even if larger sample sizes were to be assessed.

Related to this comment, the advantages and disadvantages of a between-subject versus within-subject design could also be discussed; since Cyberball and the TSST are conceptually and practically two very

different tests, a within subject design could be envisioned where subjects are exposed to both the TSST and Cyberball in a counterbalanced manner, perhaps with one week difference. The net investment in time and effort would probably have been very similar, but the added advantage would have been to compare results in cortisol and progesterone within the individual (which might not have made a difference in case of insignificant stress responses to Cyberball, but still). This is not really a critique but rather a comment that could be entered into the discussion, to alert the reader to the possibility of alternative experimental designs.

Another comment relates to the menstrual cycle phase, which the authors mentioned they couldn't control for in the current study. This is if course suboptimal and the authors already acknowledge this in the limitation section; however, they argue that the damage is minimal as they were interested in the change from baseline, rather than baseline differences, thus the chosen measure (AUCi) can compensate for the lack of menstrual cycle phase control. I am not convinced that this approach really resolves this particular problem - what if the magnitude of the change depends on the baseline level? For example, if you are already high at baseline, you won't see a strong response anymore? Even though there is no evidence for this in the current study, I would not be comfortable with accepting the advise that it is OK to not control for menstrual cycle phase when measuring progesterone levels in cycling young women.

One other aspect of the experimental design caught my attention - the authors mention that they used the essay writing about the dream job (which is part of the TSST control condition) to fill up the time difference between the Cyberball and the TSST sessions. While it is commendable to control for total time exposure, the authors have now in reality confounded the Cyberball with at least part of the TSST control condition - might that have induced some interaction effects? From Figure 2, it appears that the TSST control (TSST-C) is actually leading to a less pronounced circadian decline overall, and that the cortisol increase between sample 3 and 4 in the group of women is actually strongest in the TSST-C group. So by itself, the TSST-C might have some effects on at least cortisol; combining it with the other stress task you want to compare against might then present a suboptimal approach, and should be critically discussed as well.

On other point: In the past, effects from the biological sex of the TSST panel on the magnitude of the cortisol stress response in the test subjects has been observed. The paper doesn't mention the sex of the judges - were those mixed (one man, one woman), or unisex (which?), or changing depending on availability? For the Cyberball, what was the sex of the other 'players'? Depending on the setup, this should either be evaluated, or mentioned as a possible limitation as it could explain additional variation in the endocrine data.

Finally, one technical question: What was the reason for the three subsequent freeze-thaw cycles prior to performing the assay? Was that a recommendation by the manufacturer? This information should be added to the methods, rather than just stating it.

Overall, I think that this study is an important contribution to the literature. The observation that Cyberball might not cause an increase in cortisol is important for many stress researchers contemplating various experimental designs but for that reason the addition of a power calculation is essential.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 16 Oct 2014

Michelle Wirth, University of Notre Dame, USA

1. We agree that a power analysis is necessary in order to have confidence in these null findings. We performed a post-hoc power analysis, calculated using a partial eta squared of .10 (the small effect size we obtained in our Group X Sex ANOVA on cortisol), and found that, with our sample size, we had power of .90 to detect this effect. Therefore, we believe we had adequate power to detect even quite small effects.

2. Within-subject designs: although this was a possibility, there would have been several disadvantages to performing this study using a within-subjects design, including habituation and order effects. It is well-known that participants quickly habituate in their cortisol responses to the TSST. Laboratory stressors might show cross-habituation, so that null effects in subsequent sessions would be ambiguous: are they due to habituation, or because those conditions do not produce hormonal responses? In general, order effects are common in within-subject hormone research (e.g. Wirth, Scherer, Hoks & Abercrombie, 2011, *Psychoneuroendocrinology*; Herzmann, Young, Bird & Curran, 2012, *Brain Research*). With four conditions (since the TSST and Cyberball have their own, separate control conditions) we would have to potentially consider the effects of 16 different orders of conditions- which is prohibitive logistically and also would have reduced our power. This is why for this study, we administered conditions between subjects. Within-subject designs are certainly a viable option for other kinds of studies, but did not seem practical or appropriate in this case.

3. Menstrual phase: we agree that the magnitude of change in a hormone could depend on baseline levels. This is why we did conduct our analyses with menstrual phase ("days ago" variable) as a covariate. See third paragraph of Data Analysis: "self-reported days since period, entered as a covariate, did not moderate the effect of Group on either cortisol AUC_i or progesterone AUC_i in women." This test would show us if the magnitude of the change in progesterone, as represented by AUC_i, was affected by baseline progesterone levels, as represented by days since period. However, we agree that this is a suboptimal approach, since days since last period is a suboptimal method of assessing menstrual phase. We agree that future research assessing progesterone levels in women should more carefully control for menstrual phase. We have added a recommendation to this effect in the Limitations, under Discussion.

4. Writing task as potential confound: first, in case there is any confusion, an important difference between the writing task in the Trier control and Cyberball versus the Trier stress condition is that, in the Trier stress condition, participants knew that they are taking notes for a speech they would have to give. In the other two conditions, participants were told that the essay's content would not be judged or evaluated in any way. As for overlap between the Trier control and the Cyberball conditions, writing an essay under these instructions has been used in Trier control conditions frequently in our and other laboratories (e.g. our colleague Jessica Payne; see also Het *et al.* 2009, *Psychoneuroendocrinology*). Having to write about one's dream job when one is told the writing will not be evaluated or judged does not seem to elicit any cortisol response. Therefore, in the Cyberball condition, we expected that any response would be due to the Cyberball rejection manipulation, not due to the writing task. In figure 2C, the greatest increase between time point 3 and time point 4 in women is actually in the Trier stress group (dotted line with black squares is the Trier control

group), although the error bars are overlapping for these two groups such that there are no significant differences.

5. The panel of judges always consisted of one man and one woman; we agree that this is important information and have added this to the Methods, under Tasks. In Cyberball, the other "players" were always both of the same sex as the participant. This was already mentioned in passing under Tasks, however we have now made it more explicit.

6. Freeze-thaw cycles are part of our standard processing steps for saliva samples in order to break up long-chain mucopolysaccharides, and thereby make the saliva less viscous and able to be pipetted accurately. See e.g. Wirth and Schultheiss 2006, *Hormones and Behavior*. This information has been added to the hormones section of the Methods.

Competing Interests: No competing interests to report.

Referee Report 29 September 2014

doi:10.5256/f1000research.5482.r6019

Richard Slatcher
Department of Psychology, Wayne State University, Detroit, MI, USA

I enjoyed reading this paper and think that it will make an important contribution to the literature. This is one of those instances in which a null finding can potentially be very informative in terms of theory development. Although there are many different possibilities of why Gaffey & Wirth did not find effects of rejection on cortisol, the theory that they put forth—that rejection does not necessitate high energy mobilization needs compared to the acute stress of giving a speech—is an interesting and compelling one. Those of us who study cortisol often mention the energy mobilization aspects of cortisol as almost an afterthought. This paper puts the idea of energy mobilization front and center. Self-evaluative threat is clearly a key part of why the Trier Social Stress Test (TSST) typically "works" to raise cortisol levels. This paper suggests that social evaluative threat itself may be not sufficient for triggering a cortisol response, but rather that high energy mobilization may be necessary as well.

Although I generally really liked this paper and thought that it was well written, I do have some concerns and suggestions, which I hope that the authors will find useful in revising this paper.

1. In the 'Introduction', it is said that high levels of glucocorticoids have been proposed as one mechanism mediating the links between isolation and poor health. I think that perhaps the term "cortisol dysregulation" or "HPA axis dysregulation" might be better than "high levels." The evidence that cortisol responses during the TSST are associated with poorer health is very slim, as are the data linking total cortisol output over the day (both AUCg and AUCi) to physical health. There is some evidence that flatter diurnal cortisol slopes (a less steep decline in cortisol across the day) are associated with poorer physical health/mortality, but the picture is clearly incredibly complex, potentially involving glucocorticoid resistance and many other intervening factors. The idea that higher cortisol is bad is, I think, overly simplistic; "dysregulation" would more accurate.

2. Is loneliness the same as rejection? These two terms are used almost interchangeably in the Introduction, but are quite different. One can feel lonely without being rejected. Being rejected can

lead to feelings of loneliness, but so can many other social conditions (not living near people, not going out much, being shy, etc.). I think it is important to make distinctions between loneliness and rejection/ostracism in the paper. Also, to what extent might chronic loneliness drive HPA activity compared to fleeting loneliness? Is it perhaps only chronic loneliness that really matters? Adam *et al.*'s 2006 *PNAS* paper looking at day-to-day loneliness and cortisol might be informative here.

3. My guess is that most readers will be less familiar with the progesterone literature than the cortisol literature, so I think a little bit more thorough description of progesterone and its health and psychological correlates would be helpful. It is said in the 'Introduction' that there are important implications of progesterone "for understanding psychological disorders." A couple of specific implications would be helpful there. In the paragraph following that one, it was quite unclear to me whether elevated progesterone is good or bad for one's health (or neither, or unknown). For example, the statement "links progesterone to the beneficial effects of helping behavior on cardiovascular recovery from stress" makes it sounds like higher levels of progesterone are good. However, here and elsewhere in the paper, the direction of association (positive or negative) is omitted. Please go through and add in the direction of association throughout the paper, just so the reader is clear about which way the direction of effects are going. Another example "progesterone and not cortisol showed (positive?) associations with affiliation."

4. Were *a priori* power analyses conducted? If so, please report. If not, please report achieved power.

5. Was self-reported loneliness and the subjective feeling of being rejected assessed after Cyberball? "I felt non-existent during the game" seems related to feeling rejected, but it is not the same thing. A concern that I had was whether or not Cyberball is a powerful enough rejection manipulation to elicit a cortisol response in the lab. To confidently refute the hypothesis that rejection leads to cortisol increases, I think we need to first know that the participants felt strongly rejected and/or lonely afterwards. I realize that Cyberball is a widely used paradigm, but it still may not be powerful enough to effectively test the rejection-cortisol association *if* people are not feeling powerfully rejected. Related to this, it is said in 'Cyberball manipulation check' that "participants were clearly aware of the exclusion and had negative feelings about it." Clearly they were aware of the exclusion, but the results presented do not indicate that participants felt bad about it. Please rephrase/clarify.

6. I think a great point of discussion could center around the idea of energy mobilization. If energy mobilization is key to why the TSST raises cortisol and why Cyberball doesn't, then how might energy mobilization be directly tested? As I said at the beginning of my review, this is an exciting and provocative idea that could potentially be directly tested. How might one go about doing that? Looking at pre- and post-task depletion and/or fatigue? Changes in blood glucose? Some more discussion of this idea and steps forward would be a nice addition to the paper.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 16 Oct 2014

Michelle Wirth, University of Notre Dame, USA

1. We agree completely. We have made changes to the second paragraph of the Introduction to reflect the fact that all kinds of dysregulation in the HPA axis – including, as you mention, flattened/blunted diurnal patterns of cortisol – have been associated with poorer health. The changes we have made are subtle, however, because we stand by the research cited which found associations between (a) loneliness with higher cortisol, (b) social isolation with robust glucocorticoid increases in laboratory animals, and (c) chronically high cortisol with poorer health outcomes (not to say that is the only type of HPA dysregulation associated with health issues). We acknowledge at the end of this paragraph that these relationships are complex, and poorly understood, which I think also underscores your point.

2. Loneliness vs. rejection: This is an excellent point, and though we do not believe these two conditions are identical or interchangeable, we do believe they involve similar kinds of subjective feelings, and also may share some physiological consequences (i.e., HPA axis activation). We have rewritten the second paragraph of the Introduction to be clearer that we are not conflating the two concepts, but that they are associated. We have also hinted at the need to study both acute rejection and long-term/chronic rejection or other forms of social isolation, although this is touched on more directly in the Discussion, where we are citing Adam *et al.* (2006), along with the possibility that HPA axis effects might only be seen in chronic rejection or isolation rather than fleeting feelings in laboratory studies. I hesitate to discuss the Adam *et al.* paper in more depth here, since those authors were examining diurnal cortisol profiles and cortisol awakening response, which is not necessarily comparable to response to a laboratory stressor. However, we do note that Adam *et al.* found that subjective loneliness associated with next-day cortisol – not same-day – suggesting longer-term effects. Unfortunately, we do not have next-day cortisol awakening response data in the current study.

3. Progesterone / detail and directions of associations with health: Excellent point, and we would argue that to some extent, your point in #1 applies here as well, that it is likely dysregulation in progesterone / allopregnanolone rather than overall higher or lower levels that are associated with health issues. That said, there is evidence that psychological disorders as diverse as depression, PTSD, and schizophrenia are associated with lower levels of allopregnanolone compared with healthy controls (see Wirth 2011, *Frontiers in Endocrinology*). Progesterone, as the precursor for allopregnanolone, does not always show such differences; some evidence points to dysregulation in the enzymes needed to produce allopregnanolone from progesterone in psychopathologies, rather than differences in progesterone levels. However, measuring progesterone changes in response to social rejection and other stressors is relevant since progesterone levels are one factor (along with enzyme availability) that helps determine levels of neurosteroids such as allopregnanolone. As requested, we have added an example in the introduction (lower allopregnanolone in depression), and we have clarified the association between progesterone and affiliation in the sentence noted.

4. Unfortunately, we did not conduct *a priori* power analyses. We do now report achieved power, however. As per our response to Jens Pruessner, the other reviewer: We performed a post-hoc power analysis, calculated using the partial eta squared obtained in our omnibus ANOVA of .10, and found that, with our sample size of 131, we had power of .90 to detect an effect of this size. Therefore, we believe we had adequate power to have detected even relatively small effects.

5. We unfortunately do not have data specifically on how rejected participants felt after Cyberball. The question appeared ambiguously on the questionnaire we administered: It was worded “How accepted/rejected did you feel?”, without explicit instructions to circle or indicate “accepted” vs. “rejected” – which most participants did not do, making their numerically rated responses uninterpretable. However, we would argue that from the post-Cyberball “manipulation check” questionnaire, participants not only were aware of the exclusion, but also had negative feelings. For example, as stated in the Results, participants in the Cyberball exclusion condition rated themselves as feeling more *like an outsider*, more *nonexistent*, and *less in control*, compared to the Cyberball inclusion condition. These are arguably negative feelings, which were increased in the exclusion condition. Exclusion condition participants also scored lower on ratings of the other game-players perceiving them as worthy and likable people, which, while not a direct report of feelings, can reasonably be expected to be associated with feelings of rejection. However, it is true (and we state in the results) that there was no difference in particpants’ feelings of frustration or anger. We agree that Cyberball does not seem like a very powerful manipulation, that there may exist other social rejection manipulations might elicit more powerful feelings of rejection, and possibly also changes in cortisol or progesterone. We address this in the Discussion.

6. Directly testing energy utilization is an excellent idea – however, to accomplish this in practice could be logistically quite difficult. To accurately measure energy usage, participants are typically tested for several hours in sealed chambers so that all their oxygen consumption and carbon dioxide output can be measured (i.e. calorimetry). Overnight stays are typically used to calculate basal metabolic rate, as for accurate measurement the sympathetic nervous system must not be activated due to novelty of the test environment, etc. let alone laboratory stressors. I don’t think blood glucose measurements would be an accurate way to measure energy consumption during stressors, since as glucose is depleted for energy during the stressor, it will be rapidly replenished by glucocorticoids, making interpretation of changes difficult. There are kinesiology and exercise physiology laboratories studying energy usage in humans, so one possible step forward might be a collaboration between e.g. kinesiologists and psychologists. However, one barrier is that these two groups are interested in very different research questions. For these reasons, directly measuring energy utilization is perhaps not an ideal first step forward for this line of research.

Competing Interests: No competing interests to report.