ACADEMIC SENATE
Committee on Academic Planning and Review

CALIFORNIA STATE UNIVERSITY EAST BAY

ASSESSMENT REPORT

College	Science
Department	Psychology
Program	Psychology BA/BS
Reporting for Academic Year	2024-2025
Last 5-Year Review	2023-2024
Next 5-Year Review	2028-2029
Department Chair	Murray Horne
Date Submitted	7/25/25

I. <u>SUMMARY OF ASSESSMENT</u> *(suggested length of 1-2 pages)*

A. Program Learning Outcomes (PLO)

List all your PLO in this box. Indicate for each PLO its alignment with one or more institutional learning outcomes (ILO). For example: "PLO 1. Apply advanced computer science theory to computation problems (ILO 2 & 6)."

New Program Learning Outcomes as of October 5, 2024:

1. Understand key concepts and theories of psychology's content domains and identify connections to personal and social issues
2. Apply scientific and statistical reasoning to interpret and conduct research on psychological phenomena with openness and transparency (ILO 1: Critical Thinking).
3. Evaluate the ethics of psychological science.
4. Demonstrate effective communication skills (ILO 2: Written Communication).
5. Describe academic and career options within psychology.
6. Recognize how Justice, Equity, Diversity, and Inclusion objectives are advanced within psychological science.

B. Program Learning Outcome(S) Assessed

List the PLO(s) assessed. Provide a brief background on your program's history of assessing the PLO(s) (e.g., annually, first time, part of other assessments, etc.)

During the 2016-2017 and 2021-2022 school years, we assessed PLO 4 using the CSUEB ILO Written Communication Rubric. During the 2017-2018, 2018-2019, and 2023-2024 school years, we created and revised an online multiple-choice test to evaluate PLOs 1, 2, and 3. During the 2019-2020 school year, we used an empirical article analysis assignment to evaluate PLO 2. During the 2020-2021 school year, we used the online multiple-choice test to assess PLO 3 and surveyed faculty on how they cover ethics in their classes. During the 2022-2023 year, we assessed PLO 5 using a new survey that we designed to assess career knowledge and students' beliefs about skills they acquired throughout the program. This year, we revised the article analysis assignment from 2019-2020 to assess PLO 2 (and ILO 1).

C. Summary of Assessment Process

Summarize your assessment process briefly using the following sub-headings.

Instrument(s): *(include if new or old instrument, how developed, description of content)*

Students read two short empirical articles and answered ten questions that assessed their critical thinking ability. First, we evaluated students' understanding of the article content by asking questions about the theory underlying the research projects, the operationalization of variables in the studies, the researchers' hypotheses and rationale, and the results. Next, we evaluated whether students could recognize pros and cons to the researchers' approaches and identify possible alternative explanations for the results. Lastly, we asked students to design a follow-up study that operationalized the key variable in a different way.

The assessment committee (Drs. Baranski, Layous, and Miranda) evaluated student responses (each student evaluated by two raters) on the Institutional Learning Outcome Critical Thinking Rubric (approved by Academic Senate March 2016) and on more specific questions of interest to the psychology department (e.g., How well did the student explain the importance of replication?).

See Appendix A for the assignment, model answers, and the articles themselves. See Appendix B for the scoring form (completed by raters on Qualtrics)

In our 2019-2020 assessment report (also focused on PLO 2 and ILO 1 like the current assessment), the Psychology Department's assessment committee recommended that we change the evaluation to include two different articles with conflicting conclusions to provide students with a clearer opportunity to compare results and state their position. We have some evidence that this assessment allowed our students to better showcase their reasoning.

Sampling Procedure:

Our sample came from Dr. Baranski's PSYC 491E section ($n = 20$), Dr. Fencsik's PSYC 491C sections ($n = 9$), and Dr. Anguas-Wong's PSYC 200 section ($n = 45$). Dr. Baranski's students completed the assessment in class and Dr. Fencsik's and Dr. Anguas-Wong's students completed the assessment on their own time for extra credit. Participants were assured non-completion would not count against them and that their assignments would be anonymized when evaluated. Participants were also notified that their responses would be evaluated by faculty in the psychology department.

Sample Characteristics:

The PSYC 491 sections represented our advanced students who should have mastered our PLO/ILO and the PSYC 200 section represented students at the beginning of their time in our department who should likely score relatively lower on this assessment. We did not collect demographic information from participants.

Data Collection: *(include when, who, and how collected)*

Dr. Baranski completed the assessment in her class on Tuesday, April 29 (during Week 14 of the Spring 2025 semester). Drs. Anguas-Wong and Fencsik introduced the assignment as an extra credit opportunity and allowed a couple weeks for students to turn it in (Dr. Anguas-Wong during Weeks 5-7 and Dr. Fencsik during Weeks 13-14 of the Spring 2025 semester). The assignment and corresponding articles were posted to Canvas. Students downloaded the assignment, typed their answers to the questions, and then re-uploaded their completed assignment to the Canvas assignment link. Drs. Anguas-Wong and Fencsik then downloaded the completed assignments and sent them to the Psychology Department's assessment committee. Prior to evaluation, Dr. Fencsik assigned numbers to each student's entry so raters (Baranski, Layous, and Miranda) did not know which assignments came from which sections.

Data Analysis:

Because each student was evaluated by two raters, we first explored whether the two raters agreed on the ratings (i.e., interrater reliability). We looked for absolute agreement (more stringent), as well as consistency agreement (allowing for the fact certain raters may have rated people higher or lower than others, but the rise and fall of the ratings were similar). Intraclass correlations indicated that raters achieved good absolute agreement on scores, so we felt comfortable averaging across raters.

Table 1 includes the means, standard errors, and percentage competent (scoring ≥ 3) for items on the ILO Critical Thinking rubric separated by PSYC 200 and PSYC 491 sections. Table 2 includes the means and standard errors for the Psychology Department's more specific questions again separated by PSYC 200 and PSYC 491 sections.

Descriptively, PSYC 491 students scored better than PSYC 200 students on all metrics except for "Alternative Viewpoints" (on the ILO rubric) and the last two psychology-focused questions about designing a new study. That said, these differences were relatively small (Cohen's *d*s between |.05| and |.39|) and therefore mostly did not rise to the level of significance given our small sample sizes (*p*s .07 to .85), with one exception: use of evidence was significantly stronger among PSYC 491 students than PSYC 200 students, $t(72) = 2.06$, $p = .04$, $d = .44$. Percentage of competent PSYC 491 students (scoring ≥ 3) was 72.4% or higher on all ILO rubric categories except "Alternative viewpoints" (37.95%, but see Limitations section about that category).

We also compared the means on the ILO rubric from our PSYC 491sections this year ($n = 29$) to our assessment of PSYC 491 sections in 2019-2020 ($n = 29$) and found that the current year outscored 2019-2020 across all six categories: Explanation of issues: $t(56) = 5.92$, $p < .001$, $d = 1.55$; Use of evidence: $t(56) = 3.35$, $p = .001$, $d = 0.88$; Context, assumptions: $t(56) = 4.39$, $p < .001$, $d = 1.15$; Alternative viewpoints: $t(56) = 1.66$, $p = .10$, $d = 0.44$; Statement of position: $t(56) = 5.60$, $p < .001$, $d = 1.47$; Conclusions, implications, and consequences: $t(56) = 5.45$, $p < .001$, $d = 1.43$. These increases could be due to our improved instrument, better instructional attention to critical thinking, or suspected AI-use (see Limitations section).

Table 1

Means (Standard Errors) and Percentage Competent on ILO Critical Thinking Rubric

	PSYC 200		PSYC 491	
	M (SE)	% Comp	*M (SE)*	% Comp
Explanation of Issues	3.22 (0.15)	66.7%	3.48 (0.13)	82.8%
Use of Evidence	3.14 (0.16)	68.9%	3.53 (0.11)	89.7%
Context, assumptions	3.11 (0.16)	62.2%	3.25 (0.13)	72.4%
Alternative viewpoints	2.72 (0.17)	53.3%	2.54 (0.19)	37.9%
Statement of position	3.16 (0.15)	68.9%	3.27 (0.15)	72.4%
Conclusions, implications, and consequences	3.21 (0.14)	31.1%	3.25 (0.16)	75.9%

Note. 1-4 scale, with 4 being the most competent. *n* = 45 for PSYC 200 and *n* = 29 for PSYC 491. % Comp = Percentage Competent. If average scores were 3 or above, students were deemed competent.

Table 2

Means (Standard Errors) for Focused Questions

	PSYC 200	PSYC 491
How well did the student explain the main topic?	3.99 (0.19)	4.41 (0.13)
How well did the student explain the importance of the main topic?	3.84 (0.20)	4.02 (0.22)
How well did the student explain the importance of replication?	3.97 (0.18)	4.13 (0.22)
How well did the students describe the hypotheses from Carney et al. (2010)	4.12 (0.19)	4.34 (0.22)
How well did the student describe the operationalization of power posing?	3.62 (0.19)	3.79 (0.18)
Did the student accurately summarize the results from Carney et al. (2010)?	4.04 (0.17)	4.45 (0.21)
Did the student accurately assess whether Ranehill et al. (2015) replicated Carney et al. (2010)?	3.77 (0.18)	4.23 (0.23)
Did the student accurately assess whether the results supported the authors' hypotheses?	3.14 (0.21)	3.59 (0.26)
How well did the students design a new study to test power posing?	3.43 (0.20)	3.23 (0.26)
To what degree did their study address the potential limitations of Carney et al. (2010) and Ranehill et al. (2015)?	3.31 (0.21)	3.04 (0.27)

Note. These questions were on a 5-point scale from 1 = *Not at all* to 5 = *Extremely*.

D. Summary of Assessment Results

Summarize your assessment results briefly using the following sub-headings.

Main Findings:

Our advanced students (PSYC 491) outperformed our PSYC 200 on almost all evaluated categories and the percentage of competent PSYC 491 students was above 70% in 5/6 ILO Critical Thinking rubric categories. Our current PSYC 491 students also outperformed their counterparts from the 2019-2020 critical thinking assessment.

Recommendations for Program Improvement: *(changes in course content, course sequence, student advising)*

Although these scores demonstrate competence among most of our advanced students across most critical thinking categories, we still see room for improvement in considering alternative viewpoints and designing a new study to address the shortcomings of the reviewed studies. Our suggestion for program improvement is to encourage instructors across the curriculum to provide opportunities for students to critique the research they read and design studies that might address shortcomings in current literature.

Limitations:

Although we think this instrument was an improvement over the 2019-2020 instrument, we should further improve the “Alternative viewpoints” question. Students were still able to say that they did not have an alternative viewpoint and therefore scored relatively lower in that area. We likely need to state that they *need* to come up with an alternative explanation (instead of “Can you think of any alternative explanations…?” to which they often say no).

Additionally, raters suspected AI use in some of the completed assignments. Coders rated the degree to which they thought the student used AI to complete the assignment: 1 = *Not at all (e.g., no AI was used to generate responses)*, 2 = *Somewhat (e.g., for editing purposes only)*, 3 = *Quite a bit (e.g., idea generation, some response writing*), 4 = *Completely* (e.g., *all responses were AI-generated)*. 35.5% of PSYC 200 students, 10% of Dr. Baranski’s PSYC 491 section, and 33.3% of Dr. Fencsik’s PSYC 491 sections had an average score of 2.5 or above on this scale. Importantly, the sections with the higher suspected AI use completed the assignment on their own time whereas the section with the lower suspected AI use completed the assignment in class. The high percentage of suspected AI use among PSYC 200 students could also account for their scores being higher than might have been expected. Also, the section with higher suspected AI use in PSYC 491 was the smaller one, so that advantage was weighted to a lower degree than the other PSYC 491 section with the lower use.

Next Step(s) for Closing the Loop: *(recommendations to address findings, how & when)*

We will discuss the results from the 2024-2025 assessment during a faculty meeting fall semester.

We discussed the results of the previous year's (2023-2024) assessment of PLO 1 at a departmental faculty meeting on August 26, 2024. We mostly discussed why our students only scored about 50% on the assessment of psychology topic areas. The faculty seemed to converge on the idea that we should not use the GRE Psychology Subject questions. Instead, it might we should ask faculty in specific subject areas to agree on topics they all cover in their courses (e.g., Physiological Psychology) and pose multiple-choice questions that address those topics. Previously, we had started with the GRE questions and then asked faculty to choose and revise, but perhaps a completely bottom-up approach might be best to accurately reflect what faculty are covering in the curriculum. Running the questions by all lecturers would also be ideal to ensure we accurately reflect how they approach the topic areas.

Additionally, the faculty recommended that we only include scores in the report from people who took that topic course and eliminate items from the assessment that had a very low percentage correct.

We also discussed that scores on the ethics and research method sections of the assessment were relatively high, perhaps because they are tackled in multiple classes and should continue to be.

Other Reflections:

In the 2019-2020 report, the assessment committee recommended that we compare PSYC 491 students to PSYC 200 students to reflect improvement in critical thinking over the course of the program, so we implemented that this year. That said, we wonder if the assignment was a bit too difficult for PSYC 200 students and perhaps that's why 35.5% turned to AI for assistance. Indeed, Dr. Baranski said it took many of her PSYC 491 students the entire class period (1 hour and 40 minutes) to complete the assessment, reflecting the demanding nature of the task. In the future, perhaps we should not give this potentially discouraging assignment to PSYC 200 and instead just evaluate the percentage competent among our PSYC 491 students. Additionally, we might consider evaluating critical thinking in a less demanding way or using a pre-existing course assignment.

The suspected AI use was concerning. In future written assessments, we should likely require that students complete in class to standardize across sections and turn on the AI checker in Canvas for an official diagnostic.

E. Assessment Plans for Next Year

Summarize your assessment plans for the next year, including the PLO(s) you plan to assess, any revisions to the program assessment plan presented in your last five-year plan self-study, and any other relevant information.

During spring semester of 2026, we plan to evaluate PLO 3 (ethics) and our newly approved PLO 6 (about justice, equity, diversity, and inclusion). We will likely evaluate ethics similarly to how we have in the past, with multiple-choice questions for students and a survey of faculty asking what topics they cover and how. For PLO 6, we will likely combine with the ethics assessment and thus include some multiple-choice questions and a survey of faculty.

Article Analysis Assignment

Instructions:

CSUEB conducts periodic reviews of student work to improve our psychology curriculum. This review helps us assess students' **critical thinking skills**, particularly their ability to **interpret scientific findings and design studies in psychology**. Faculty from the Psychology Department and other departments will evaluate your responses **anonymously**. Your work will **not affect your grade**, but we ask that you complete this assignment carefully to best represent your skills.

Steps to Complete the Assignment:

1. **Download and carefully read the two assigned articles:**
   - **Carney et al. (2010)** – The original study on power posing.
   - **Ranehill et al. (2015)** – A replication attempt of Carney et al. (2010).
2. **Answer the Ariticle Analysis Questions below.**
3. **Upload your completed document to Canvas.**
   - Do **not** include your name inside the document (your instructor will know you completed it based on your Canvas submission).

Article Analysis Questions:

1. What is the main topic of these articles? How does the article by Carney et al. (2010) discuss why this topic is important to study?

The topic both studies are exploring is power posing and whether posing in a particularly powerful stance can cause someone to feel powerful. In Carney et al. (2010), authors explain that power is related to access to resources, motivation toward action, and enhanced cognitive function.

2. What is the purpose of the Ranehill et al. (2015) article? Why do you think it is important to conduct studies like Ranehill et al (2015)?

The purpose of Ranehill et al. (2015) was to conduct a conceptual replication of Carney et al. (2010) to assess the robustness and reliability of the original findings. Their study was designed to use a much larger sample (N = 200 vs. N = 42) and a blinded experimenter to eliminate potential experimenter bias.

Replication studies like Ranehill et al. (2015) are crucial in psychology because they help determine whether an effect is statistically reliable or if the original findings were false positives due to small sample sizes, methodological biases, or statistical anomalies. This is particularly important in light of the replication crisis in psychology, where many high-profile findings have failed to replicate.

Commented [1]: ILO 1: Explanation of issues

Commented [2]: ILO 1: Explanation of issues

Commented [3R2]: PLO: Apply scientific and statistical reasoning to interpret and conduct research on psychological phenomena with openness and transparency

3. What were the researchers' hypotheses in Carney et al. (2010)?

Carney et al. (2010) hypothesized that adopting high-power poses (e.g., expansive, open postures) would lead to:

- Increased feelings of power and dominance (self-report measure).
- Higher testosterone levels (associated with dominance and social status).
- Lower cortisol levels (indicating reduced stress response).
- Greater risk-taking behavior (as seen in a gambling task).

4. What was the rationale for the researcher's hypotheses? Put another way, this question asks you to explain why the researchers predicted what they predicted.

The authors thought that because testosterone is associated with status and dominance in both humans and animals, they would expect this hormone to increase after individuals embody dominance and power. Likewise, people high in dominance tend to have low levels of cortisol, thus power poses should predict lower levels of this hormone. Further, in explaining why power poses should lead to increased feelings of power and a focus on reward instead of risk, researchers point to previous work that demonstrates the impact certain facial expressions have on mood (e.g., head tilted up predicts pride).

5. How did the researchers operationalize power posing in each study? In other words, describe how power posing was represented in both Carney et al. (2010) and Ranehill et al. (2015).

For both studies, power posing was operationalized as standing in two dominant positions (i.e., leaning forward with hands on a table; sitting with hands on the head with legs propped up). Carney et al. (2010) had participants hold each pose for 1 minute and Ranehill et al. (2015) had them hold each pose for 3 minutes.

6. Please compare the results from both articles. In other words, what were the results of Carney et al. (2010), and did Ranehill et al. (2015) successfully replicate these findings? Also, note whether these results supported the authors' hypotheses.

In alignment with their hypotheses, Carney et al. (2010) found that relative to those in the low-power posing condition, high-power posers had higher levels of testosterone, lower levels of cortisol, and reported higher levels of feelings of power and tolerance for risk. Ranehill et al. (2015) successfully replicated the finding that those in the high power posing condition reported stronger feelings of power, however, they did not find any effect of testosterone, cortisol, or risk tolerance.

7. Evaluate the design of Carney et al. (2010) and Ranehill et al. (2015). How were they similar and how did they differ? Which study do you think was a stronger test for the impact of power-posing and why?

Ranehill et al.'s methodology closely followed Carney et al's. Participants were first randomized into low-power and high-power posing conditions. They then provided saliva samples before and after the intervention, and they completed assessments of self-reported power and risk tolerance. Ranehill et al. did differ in a few ways. First, Ranehill et al. asked participants to pose in each position for 3 minutes (relative to Carney et al.'s 1 minute). Ranehill also added two more risk propensity behavioral tasks: risk-taking in the loss domain and willingness to compete. *Ranehill also gave instructions via computer so that experimenters were blind to condition. Filler task was*

Commented [4]: PLO: Apply scientific and statistical reasoning to interpret and conduct research on psychological phenomena with openness and transparency

Commented [5]: ILO: Use of evidence

Commented [6]: ILO: Use of evidence
PLO: Apply scientific reasoning to interpret psychological phenomena and to design and conduct basic psychological research

Commented [7]: ILO: Context, assumptions

Commented [8R7]: ILO: Statement of position *different. No cover story in Ranehill like in Carney, but they did not give away any details of their hypotheses.* Finally, Ranehill et al. recruited 200 participants whereas Carney et al. recruited 42.

Students will likely (or should) recognize Ranehill et al. as the stronger study due to the additional behavioral tasks and increased sample size.

8. Can you think of any alternative explanations for the results of these two studies? In other words, do you think the authors' conclusions were appropriate or would you interpret these results differently? Explain your answer.

Commented [9]: ILO: Alternative viewpoints

Students may point to a self-fulfilling prophecy effect in Carney et al.'s study, where participants likely assumed they should feel powerful or, at the very least, positive in the high-power condition relative to the low-power condition. They may also point to experimenter effects.

For the Ranehill et al. study, students will likely point to discrepancies between the two studies as a reason why Ranehill did not replicate most of Carney's findings.

9. What are the implications of the result of Ranehill et al.'s replication attempt? What do their results say about the concept of power-posing? What about the importance of replication in psychology?

Commented [10]: ILO: Conclusions, implications, and consequences
ILO: Statement of position
PLO: Apply scientific and statistical reasoning to interpret and conduct research on psychological phenomena with openness and transparency

Ranehill et al.'s mostly failed replication underscores the importance of conducting high-powered, blinded experiments. Suggests that power posing may influence self-perception but not physiological or behavioral outcomes. Raises concerns about small-sample studies producing false positives. Emphasizes the need for pre-registration and open science practices. They might also point to the possibility that the impact of power-posing may only occur under specific conditions and that more research is needed to understand the complexity of this phenomenon.

10. Design a study that tests the effects of power posing using a different methodology. You may change the duration of poses, the context in which posing occurs (e.g., real-world vs. lab setting), the dependent variables measured, or other aspects of the experimental design. Explain why your proposed study is important and how it could address the potential limitations of Carney et al. (2010) and Ranehill et al. (2015).

Good study – explain manipulation of their IV, explain their DV, explain the context (the setting, the participants)

Research Report

Psychological Science
21(10) 1363–1368

DOI: 10.1177/0956797610383437
http://pss.sagepub.com

Power Posing: Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance

Dana R. Carney[1], Amy J.C. Cuddy[2], and Andy J. Yap[1]
[1]Columbia University and [2]Harvard University

Abstract

Humans and other animals express power through open, expansive postures, and they express powerlessness through closed, contractive postures. But can these postures actually cause power? The results of this study confirmed our prediction that posing in high-power nonverbal displays (as opposed to low-power nonverbal displays) would cause neuroendocrine and behavioral changes for both male and female participants: High-power posers experienced elevations in testosterone, decreases in cortisol, and increased feelings of power and tolerance for risk; low-power posers exhibited the opposite pattern. In short, posing in displays of power caused advantaged and adaptive psychological, physiological, and behavioral changes, and these findings suggest that embodiment extends beyond mere thinking and feeling, to physiology and subsequent behavioral choices. That a person can, by assuming two simple 1-min poses, embody power and instantly become more powerful has real-world, actionable implications.

Keywords

cortisol, embodiment, hormones, neuroendocrinology, nonverbal behavior, power, risk taking, testosterone

Received 1/20/10; Revision accepted 4/8/10

The proud peacock fans his tail feathers in pursuit of a mate. By galloping sideways, the cat manipulates an intruder's perception of her size. The chimpanzee, asserting his hierarchical rank, holds his breath until his chest bulges. The executive in the boardroom crests the table with his feet, fingers interlaced behind his neck, elbows pointing outward. Humans and other animals display power and dominance through expansive nonverbal displays, and these power poses are deeply intertwined with the evolutionary selection of what is "alpha" (Darwin, 1872/2009; de Waal, 1998).

But is power embodied? What happens when displays of power are posed? Can posed displays cause a person to feel more powerful? Do people's mental and physiological systems prepare them to be more powerful? The goal of our research was to test whether high-power poses (as opposed to low-power poses) actually produce power. To perform this test, we looked at the effects of high-power and low-power poses on some fundamental features of having power: feelings of power, elevation of the dominance hormone testosterone, lowering of the stress hormone cortisol, and an increased tolerance for risk.

Power determines greater access to resources (de Waal, 1998; Keltner, Gruenfeld, & Anderson, 2003); higher levels of agency and control over a person's own body, mind, and positive feelings (Keltner et al., 2003); and enhanced cognitive function (Smith, Jostmann, Galinsky, & van Dijk, 2008). Powerful individuals (compared with powerless individuals) demonstrate greater willingness to engage in action (Galinsky, Gruenfeld, & Magee, 2003; Keltner et al., 2003) and often show increased risk-taking behavior[1] (e.g., Anderson & Galinsky, 2006).

The neuroendocrine profiles of the powerful differentiate them from the powerless, on two key hormones—testosterone and cortisol. In humans and other animals, testosterone levels both reflect and reinforce dispositional and situational status and dominance; internal and external cues cause testosterone to rise, increasing dominant behaviors, and these behaviors can elevate testosterone even further (Archer, 2006; Mazur &

Corresponding Authors:
Dana R. Carney, Columbia University, Graduate School of Business, 717 Uris Hall, 3022 Broadway, New York, NY 10027-6902
E-mail: dcarney@columbia.edu
Amy J.C. Cuddy, Harvard Business School, Baker Library 449, Boston, MA 02163
E-mail: acuddy@hbs.edu

Booth, 1998). For example, testosterone rises in anticipation of a competition and as a result of a win, but drops following a defeat (e.g., Booth, Shelley, Mazur, Tharp, & Kittok, 1989), and these changes predict the desire to compete again (Mehta & Josephs, 2006). In short, testosterone levels, by reflecting and reinforcing dominance, are closely linked to adaptive responses to challenges.

Power is also linked to the stress hormone cortisol: Power holders show lower basal cortisol levels and lower cortisol reactivity to stressors than powerless people do, and cortisol drops as power is achieved (Abbott et al., 2003; Coe, Mendoza, & Levine, 1979; Sapolsky, Alberts, & Altmann, 1997). Although short-term and acute cortisol elevation is part of an adaptive response to challenges large (e.g., a predator) and small (e.g., waking up), the chronically elevated cortisol levels seen in low-power individuals are associated with negative health consequences, such as impaired immune functioning, hypertension, and memory loss (Sapolsky et al., 1997; Segerstrom & Miller, 2004). Low-power social groups have a higher incidence of stress-related illnesses than high-power social groups do, and this is partially attributable to chronically elevated cortisol (Cohen et al., 2006). Thus, the power holder's typical neuroendocrine profile of high testosterone coupled with low cortisol—a profile linked to such outcomes as disease resistance (Sapolsky, 2005) and leadership abilities (Mehta & Josephs, 2010)—appears to be optimally adaptive.

It is unequivocal that power is expressed through highly specific, evolved nonverbal displays. Expansive, open postures (widespread limbs and enlargement of occupied space by spreading out) project high power, whereas contractive, closed postures (limbs touching the torso and minimization of occupied space by collapsing the body inward) project low power. All of these patterns have been identified in research on actual and attributed power and its nonverbal correlates (Carney, Hall, & Smith LeBeau, 2005; Darwin, 1872/2009; de Waal, 1998; Hall, Coats, & Smith LeBeau, 2005). Although researchers know that power generates these displays, no research has investigated whether these displays generate power. Will posing these displays of power actually cause individuals to feel more powerful, focus on reward as opposed to risk, and experience increases in testosterone and decreases in cortisol?

In research on embodied cognition, some evidence suggests that bodily movements, such as facial displays, can affect emotional states. For example, unobtrusive contraction of the "smile muscle" (i.e., the zygomaticus major) increases enjoyment (Strack, Martin, Stepper, 1988), the head tilting upward induces pride (Stepper & Strack, 1993), and hunched postures (as opposed to upright postures) elicit more depressed feelings (Riskind & Gotay, 1982). Approach-oriented behaviors, such as touching, pulling, or nodding "yes," increase preference for objects, people, and persuasive messages (e.g., Briñol & Petty, 2003; Chen & Bargh, 1999; Wegner, Lane, & Dimitri, 1994), and fist clenching increases men's self-ratings on power-related traits (Schubert & Koole, 2009). However, no research has tested whether expansive power poses, in comparison with contractive power poses, cause mental, physiological, and behavioral change in a manner consistent with the effects of power. We hypothesized that high-power poses (compared with low-power poses) would cause individuals to experience elevated testosterone, decreased cortisol, increased feelings of power, and higher risk tolerance. Such findings would suggest that embodiment goes beyond cognition and emotion and could have immediate and actionable effects on physiology and behavior.

Method

Participants and overview of procedure

Forty-two participants (26 females and 16 males) were randomly assigned to the high-power-pose or low-power-pose condition. Participants believed that the study was about the science of physiological recordings and was focused on how placement of electrocardiography electrodes above and below the heart could influence data collection. Participants' bodies were posed by an experimenter into high-power or low-power poses. Each participant held two poses for 1 min each. Participants' risk taking was measured with a gambling task; feelings of power were measured with self-reports. Saliva samples, which were used to test cortisol and testosterone levels, were taken before and approximately 17 min after the power-pose manipulation.

Power poses

Poses were harvested from the nonverbal literature (e.g., Carney et al., 2005; Hall et al., 2005) and varied on the two nonverbal dimensions universally linked to power: expansiveness (i.e., taking up more space or less space) and openness (i.e., keeping limbs open or closed). The two high-power poses into which participants were configured are depicted in Figure 1, and the two low-power poses are depicted in Figure 2. To be sure that the poses chosen conveyed power appropriately, we asked 95 pretest participants to rate each pose from 1 (*very low power*) to 7 (*very high power*). High-power poses ($M = 5.39$, $SD = 0.99$) were indeed rated significantly higher on power than were low-power poses ($M = 2.41$, $SD = 0.93$), $t(94) = 21.03$, $p < .001$; $r = .99$.

To be sure that changes in neuroendocrine levels, powerful feelings, or behavior could be attributed only to the high-power or low-power attributes of the poses, we had 19 pretest participants rate the comfort, difficulty, and pain of the poses. Participants made all four poses (while wearing electrocardiography leads) and completed questionnaires after each pose. There were no differences between high-power and low-power poses on comfort, $t(16) = 0.24$, $p > .80$; difficulty, $t(16) = 0.77$, $p > .45$; or painfulness, $t(16) = –0.82$, $p > .42$.

Fig. 1. The two high-power poses used in the study. Participants in the high-power-pose condition were posed in expansive positions with open limbs.

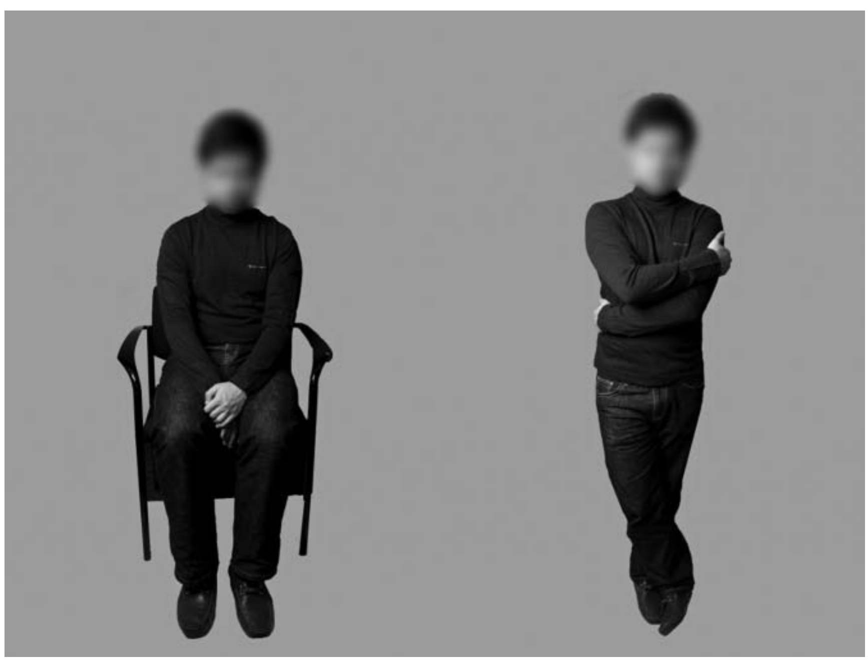

Fig. 2. The two low-power poses used in the study. Participants in the low-power-pose condition were posed in contractive positions with closed limbs.

To configure the test participants into the poses, the experimenter placed an electrocardiography lead on the back of each participant's calf and underbelly of the left arm and explained, "To test accuracy of physiological responses as a function of sensor placement relative to your heart, you are being put into a certain physical position." The experimenter then manually configured participants' bodies by lightly touching their arms and legs. As needed, the experimenter provided verbal instructions (e.g., "Keep your feet above heart level by putting them on the desk in front of you"). After manually configuring participants' bodies into the two poses, the experimenter left the room. Participants were videotaped; all participants correctly made and held either two high-power or two low-power poses for 1 min each. While making and holding the poses, participants completed a filler task that consisted of viewing and forming impressions of nine faces.

Measure of risk taking and powerful feelings

After they finished posing, participants were presented with the gambling task. They were endowed with $2 and told they could keep the money—the safe bet—or roll a die and risk losing the $2 for a payoff of $4 (a risky but rational bet; odds of winning were 50/50). Participants indicated how "powerful" and "in charge" they felt on a scale from 1 (*not at all*) to 4 (*a lot*).

Saliva collection and analysis

Testing was scheduled in the afternoon (12:00 p.m.–6:00 p.m.) to control for diurnal rhythms in hormones. Saliva samples were taken before the power-pose manipulation (approximately 10 min after arrival; Time 1) and again 17 min after the power-pose manipulation (M = 17.28 min, SD = 4.31; Time 2).

Standard salivary-hormone collection procedures were used (Dickerson & Kemeny, 2004; Schultheiss & Stanton, 2009). Before providing saliva samples, participants did not eat, drink, or brush their teeth for at least 1 hr. Participants rinsed their mouths with water and chewed a piece of sugar-free Trident Original Flavor gum for 3 min to stimulate salivation (this procedure yields the least bias compared with passive drool procedures; Dabbs, 1991). Participants provided approximately 1.5 ml of saliva through a straw into a sterile polypropylene microtubule. Samples were immediately frozen to avoid hormone degradation and to precipitate mucins. Within 2 weeks, samples were packed in dry ice and shipped for analysis to Salimetrics (State College, PA), where they were assayed in duplicate for salivary cortisol and salivary testosterone using a highly sensitive enzyme immunoassay.

For cortisol, the intra-assay coefficient of variation (CV) was 5.40% for Time 1 and 4.40% for Time 2. The average interassay CV across high and low controls for both time points was 2.74%. Cortisol levels were in the normal range at both Time 1 (M = 0.16 μg/dl, SD = 0.19) and Time 2 (M = 0.12 μg/dl, SD = 0.08). For testosterone, the intra-assay CV was 4.30% for Time 1 and 3.80% for Time 2. The average interassay CV across high and low controls for both time points was 3.80%. Testosterone levels were in the normal range at both Time 1 (M = 60.30 pg/ml, SD = 49.58) and Time 2 (M = 57.40 pg/ml, SD = 43.25). As would be suggested by appropriately taken and assayed samples (Schultheiss & Stanton, 2009), men were higher than women on testosterone at both Time 1, $F(1, 41) = 17.40$, $p < .001$, $r = .55$, and Time 2, $F(1, 41) = 22.55$, $p < .001$, $r = .60$. To control for sex differences in testosterone, we used participant's sex as a covariate in all analyses. All hormone analyses examined changes in hormones observed at Time 2, controlling for Time 1. Analyses with cortisol controlled for testosterone, and vice versa.[2]

Results

One-way analyses of variance examined the effect of power pose on postmanipulation hormones (Time 2), controlling for baseline hormones (Time 1). As hypothesized, high-power poses caused an increase in testosterone compared with low-power poses, which caused a decrease in testosterone, $F(1, 39) = 4.29$, $p < .05$; $r = .34$ (Fig. 3). Also as hypothesized, high-power poses caused a decrease in cortisol compared with low-power poses, which caused an increase in cortisol, $F(1, 38) = 7.45$, $p < .02$; $r = .43$ (Fig. 4).

Also consistent with predictions, high-power posers were more likely than low-power posers to focus on rewards—86.36% took the gambling risk (only 13.63% were risk averse). In contrast, only 60% of the low-power posers took the risk (and 40% were risk averse), $\chi^2(1, N = 42) = 3.86$, $p < .05$; $\Phi = .30$. Finally, high-power posers reported feeling significantly more "powerful" and "in charge" (M = 2.57, SD = 0.81) than low-power posers did (M = 1.83, SD = 0.81), $F(1, 41) = 9.53$, $p < .01$; $r = .44$. Thus, a simple 2-min power-pose manipulation was enough to significantly alter the physiological, mental, and feeling states of our participants. The implications of these results for everyday life are substantial.

Discussion

Our results show that posing in high-power displays (as opposed to low-power displays) causes physiological, psychological, and behavioral changes consistent with the literature on the effects of power on power holders—elevation of the dominance hormone testosterone, reduction of the stress hormone cortisol, and increases in behaviorally demonstrated risk tolerance and feelings of power.

These findings advance current understanding of embodied cognition in two important ways. First, they suggest that the effects of embodiment extend beyond emotion and cognition, to physiology and subsequent behavioral choice. For example, as described earlier, nodding the head "yes" leads a person to be more easily persuaded when listening to a persuasive appeal, and smiling increases humor responses. We suggest that these simple behaviors, a head nod or a smile, might also cause physiological changes that activate an entire trajectory of psychological, physiological, and behavioral shifts—essentially altering the course of a person's day. Second, these results suggest that any psychological construct, such as power, with a signature pattern of nonverbal correlates may be embodied.

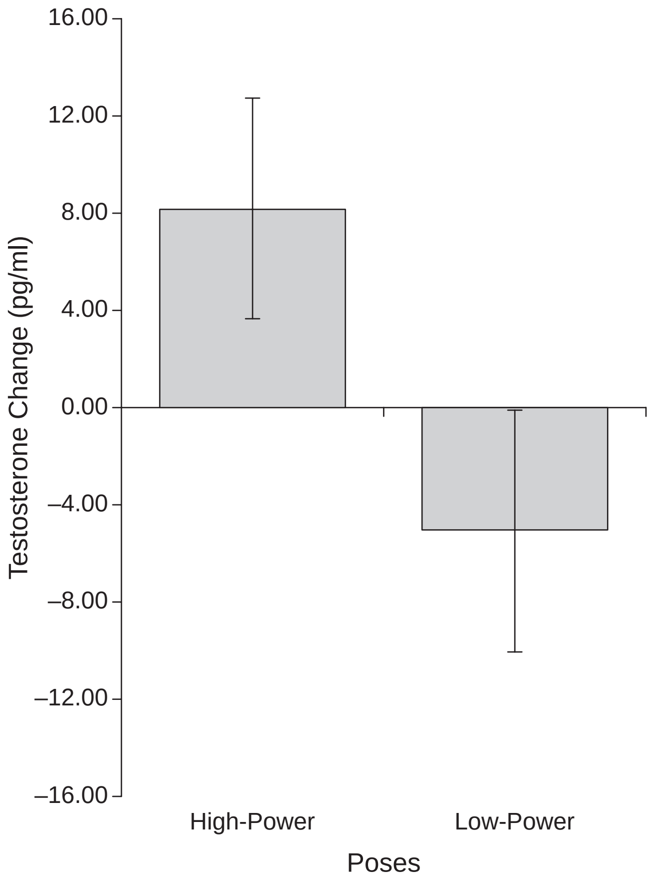

Fig. 3. Mean changes in the dominance hormone testosterone following high-power and low-power poses. Changes are depicted as difference scores (Time 2 – Time 1). Error bars represent standard errors of the mean.

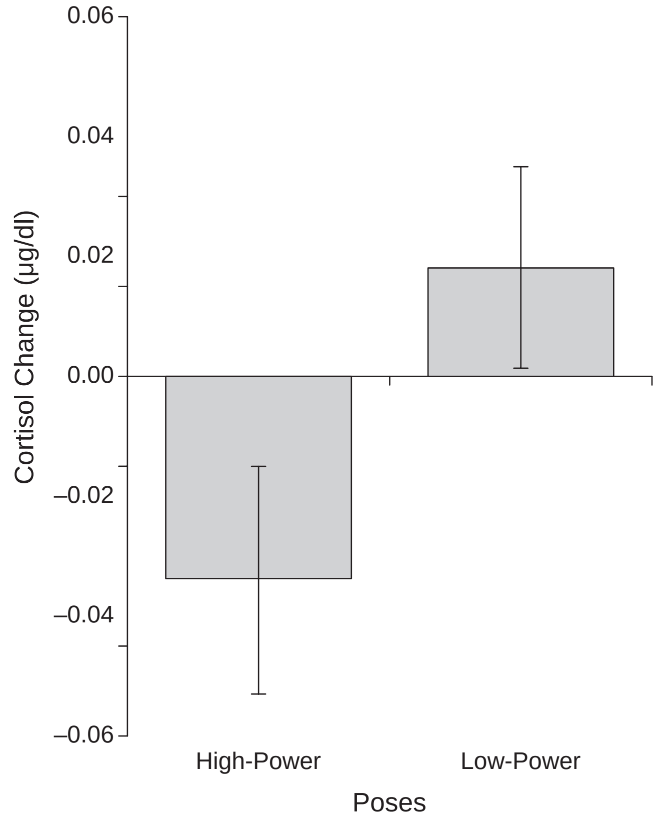

Fig. 4. Mean changes in the stress hormone cortisol following high-power and low-power poses. Changes are depicted as difference scores (Time 2 – Time 1). Error bars represent standard errors of the mean.

These results also offer a methodological advance in research on power. Many reported effects of power are limited by the methodological necessity of manipulating power in a laboratory setting (e.g., complex role assignments). The simple, elegant power-pose manipulation we employed can be taken directly into the field and used to investigate ordinary people in everyday contexts.

Is it possible that our findings are limited to the specific poses utilized in this experiment? Although the power-infusing attribute of expansiveness and the poses that capture it require further investigation, findings from an additional study (N = 49) suggest that the effects reported here are not idiosyncratic to these specific poses. In addition to the poses used in the current report, an additional three high-power poses and an additional three low-power poses produced the same effects on feelings of power, $F(1, 48) = 4.38$, $p < .05$, $r = .30$, and risk taking, $\chi^2(1, N = 49) = 4.84$, $p < .03$, $\Phi = .31$.

By simply changing physical posture, an individual prepares his or her mental and physiological systems to endure difficult and stressful situations, and perhaps to actually improve confidence and performance in situations such as interviewing for jobs, speaking in public, disagreeing with a boss, or taking potentially profitable risks. These findings suggest that, in some situations requiring power, people have the ability to "fake it 'til they make it." Over time and in aggregate, these minimal postural changes and their outcomes potentially could improve a person's general health and well-being. This potential benefit is particularly important when considering people who are or who feel chronically powerless because of lack of resources, low hierarchical rank in an organization, or membership in a low-power social group.

Acknowledgments

We are gratefully indebted to the following individuals for their insight, support, and assistance with this research: Daniel Ames, Max Bazerman, Joe Ferrero, Alan Fiske and lab, Adam Galinsky, Deborah Gruenfeld, Lucia Guillory, Brian Hall, Bob Josephs, Brian Lucas, Malia Mason, Pranj Mehta, Michael Morris, Joe Navarro, Michael Norton, Thomas Schubert, Steve Stroessner, and Bill von Hippel.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Notes

1. The effect of power on risk taking is moderated by factors such as prenatal exposure to testosterone (Ronay & von Hippel, in press).
2. Cortisol scores at both time points were sufficiently normally distributed, except for two outliers that were more than 3 standard deviations above the mean and were excluded; testosterone scores at both time points were sufficiently normally distributed, except for one outlier that was more than 3 standard deviations above the mean and was excluded.

References

Abbott, D.H., Keverne, E.B., Bercovitch, F.B., Shively, C.A., Mendoza, S.P., Saltzman, W., et al. (2003). Are subordinates always stressed? A comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior*, *43*, 67–82.

Anderson, C., & Galinsky, A.D. (2006). Power, optimism, and the proclivity for risk. *European Journal of Social Psychology*, *36*, 511–536.

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews*, *30*, 319–345.

Booth, A., Shelley, G., Mazur, A., Tharp, G., & Kittok, R. (1989). Testosterone and winning and losing in human competition. *Hormones and Behavior*, *23*, 556–571.

Briñol, P., & Petty, R.E. (2003). Overt head movements and persuasion: A self-validation analysis. *Journal of Personality and Social Psychology*, *84*, 1123–1139.

Carney, D.R., Hall, J.A., & Smith LeBeau, L. (2005). Beliefs about the nonverbal expression of social power. *Journal of Nonverbal Behavior*, *29*, 105–123.

Chen, M., & Bargh, J.A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*, 215–224.

Coe, C.L., Mendoza, S.P., & Levine, S. (1979). Social status constrains the stress response in the squirrel monkey. *Physiology & Behavior*, *23*, 633–638.

Cohen, S., Schwartz, J.E., Epel, E., Kirschbaum, C., Sidney, S., & Seeman, T. (2006). Socioeconomic status, race, and diurnal cortisol decline in the Coronary Artery Risk Development in Young Adults (CARDIA) study. *Psychosomatic Medicine*, *68*, 41–50.

Dabbs, J.M. (1991). Salivary testosterone measurements: Collecting, storing, and mailing saliva samples. *Physiology & Behavior*, *42*, 815–817.

Darwin, C. (2009). *The expression of the emotions in man and animals*. New York, NY: Oxford. (Original work published 1872)

de Waal, F. (1998). *Chimpanzee politics: Power and sex among apes*. Baltimore, MD: Johns Hopkins University Press.

Dickerson, S.S., & Kemeny, M.E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*, 355–391.

Galinsky, A.D., Gruenfeld, D.H., & Magee, J.C. (2003). From power to action. *Journal of Personality and Social Psychology*, *87*, 327–339.

Hall, J.A., Coats, E.J., & Smith LeBeau, L. (2005). Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin*, *131*, 898–924.

Keltner, D., Gruenfeld, D.H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*, 265–284.

Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral & Brain Sciences*, *21*, 353–397.

Mehta, P.H., & Josephs, R.A. (2006). Testosterone change after losing predicts the decision to compete again. *Hormones and Behavior*, *50*, 684–692.

Mehta, P.H., & Josephs, R.A. (2010). *Dual-hormone regulation of dominance*. Manuscript in preparation.

Riskind, J.H., & Gotay, C.C. (1982). Physical posture: Could it have regulatory or feedback effects on motivation and emotion? *Motivation and Emotion*, *6*, 273–298.

Ronay, R., & von Hippel, W. (in press). Power, testosterone and risk-taking: The moderating influence of testosterone and executive functions. *Journal of Behavioral Decision Making*.

Sapolsky, R.M. (2005). The influence of social hierarchy on primate health. *Science*, *308*, 648–652.

Sapolsky, R.M., Alberts, S.C., & Altmann, J. (1997). Hypercortisolism associated with social subordinance or social isolation among wild baboons. *Archives of General Psychiatry*, *54*, 1137–1143.

Schubert, T.W., & Koole, S.L. (2009). The embodied self: Making a fist enhances men's power-related self-conceptions. *Journal of Experimental Social Psychology*, *45*, 828–834.

Schultheiss, O.C., & Stanton, S.J. (2009). Assessment of salivary hormones. In E. Harmon-Jones & J.S. Beer (Eds.), *Methods in the neurobiology of social and personality psychology* (pp. 17–44). New York, NY: Guilford.

Segerstrom, S., & Miller, G. (2004). Psychological stress and the human immune system: A meta-analytic study of 30 years of inquiry. *Psychological Bulletin*, *130*, 601–630.

Smith, P.K., Jostmann, N.B., Galinsky, A.D., & van Dijk, W.W. (2008). Lacking power impairs executive functions. *Psychological Science*, *19*, 441–447.

Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, *64*, 211–220.

Strack, F., Martin, L.L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*, 768–777.

Wegner, D.M., Lane, J.D., & Dimitri, S. (1994). The allure of secret relationships. *Journal of Personality and Social Psychology*, *66*, 287–300.

Commentary

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2015, Vol. 26(5) 653–656

DOI: 10.1177/0956797614553946
pss.sagepub.com
SAGE

Assessing the Robustness of Power Posing: No Effect on Hormones and Risk Tolerance in a Large Sample of Men and Women

Eva Ranehill[1], Anna Dreber[2], Magnus Johannesson[2], Susanne Leiberg[1], Sunhae Sul[3], and Roberto A. Weber[1]
[1]Department of Economics, University of Zurich; [2]Department of Economics, Stockholm School of Economics; and [3]Department of Psychological and Brain Sciences, Dartmouth College

Received 4/24/14; Revision accepted 8/19/14

In a growing body of research, psychologists have studied how physical expression influences psychological processes (see Riskind & Gotay, 1982; Stepper & Strack, 1993, for early contributions to this literature). A recent strand of literature within this field has focused on how physical postures that express power and dominance (*power poses*) influence psychological and physiological processes, as well as decision making (e.g., Carney, Cuddy, & Yap, 2010; Cesario & McDonald, 2013; Yap, Wazlawek, Lucas, Cuddy, & Carney, 2013). Carney et al. found that power posing affected levels of hormones such as testosterone and cortisol, financial risk taking, and self-reported feelings of power in a sample of 42 participants (randomly assigned to hold poses suggesting either high or low power). We conducted a conceptual replication study with a similar methodology as that employed by Carney et al. but using a substantially larger sample ($N = 200$) and a design in which the experimenter was blind to condition. Our statistical power to detect an effect of the magnitude reported by Carney et al. was more than 95% (see the Supplemental Material available online). In addition to the three outcome measures that Carney et al. used, we also studied two more behavioral tasks (risk taking in the loss domain and willingness to compete).

Consistent with the findings of Carney et al., our results showed a significant effect of power posing on self-reported feelings of power. However, we found no significant effect of power posing on hormonal levels or in any of the three behavioral tasks.

Method

An initial power analysis based on the effect sizes in Carney et al. (power = 0.8, $\alpha = .05$) indicated that a sample size of 100 participants would be suitable. On the basis of the results of these first 100 observations, we decided to collect data from another 100 participants to further increase the reliability of our results. Of the final sample of 200, 98 were women and 102 men.

Our design closely followed Carney et al.'s (full details of our methodology, including the instructions given to participants in all tasks, can be found in the Supplemental Material). Each session involved a single participant who was randomly assigned to a high-power- or low-power-pose condition. In each condition, participants first provided a saliva sample. They then adopted two body positions, taken from those in Carney et al., while performing a filler task. Next, participants completed a risk task in both the gain domain, as in Carney et al., and the loss domain. In the gain domain, participants had to choose between winning a predetermined amount of money and a gamble that had an equal probability of resulting in a greater win or no win at all. The loss domain worked similarly, except that participants chose between losing a predetermined amount and a gamble that would result in either a greater loss or no loss. Each risk task consisted of six binary choices between the safe and the risky option (Carney et al. used one such choice in the gain domain only). Task order was fixed to keep the results in the risk task in the gain domain comparable with those of Carney et al. Our measure of risk tolerance was the proportion of risky choices made in each domain (using only the single decision directly comparable with the one in Carney et al. did not change our results).

Corresponding Author:
Eva Ranehill, University of Zurich, Department of Economics, Blümlisalpstrasse 10 Zurich 8006, Switzerland
E-mails: eva.ranehill@econ.uzh.ch, eva.ranehill@gmail.com

After the risk task, we measured competitiveness by asking participants to choose whether to solve math exercises under a competitive tournament-style payment scheme or a noncompetitive payment scheme (Niederle & Vesterlund, 2007). Willingness to compete was measured from this binary choice. As in Carney et al., the behavioral tasks were incentivized (see the Supplemental Material). Participants did not receive feedback on the outcomes of the behavioral tasks until they received their payment at the end of the study. After completing the tasks, but before being informed about their payment, participants provided a second saliva sample and filled out a postexperiment questionnaire. The questionnaire measured, among other things, self-reported feelings of power on a 4-point scale, from 1, *not at all powerful*, to 4, *very powerful* (see the Supplemental Material). As in Carney et al., saliva samples were taken before and approximately 17 min after the power-pose manipulation. The samples were collected in privacy, using passive drool procedures, and frozen immediately. Also as in Carney et al., participants were tested in the afternoon (12:00 p.m.–6:00 p.m.) and instructed not to eat, drink, or brush their teeth for the hour preceding the study (see the Supplemental Material).

A few details differed between our design and that of Carney et al. First, whereas participants in Carney et al.'s study held each position for 1 min, we extended this time to 3 min. Second, Carney et al. manually configured participants' positions before leaving the room, whereas participants in our study received instructions on a computer. As in Carney et al.'s study, participants were recorded while posing; the recordings verified that participants complied with task instructions.[1] Because the instructions were given via computer, the experimenter was blind to experimental condition, which negated potential experimenter effects. Third, the filler task in our study involved constructing words from letters and spaces; in Carney et al.'s study, the task was to form impressions of faces. Finally, we did not use deception, but informed participants briefly in the consent form that the study investigated whether physical position influences hormone levels and behavior (the instructions and consent form are in the Supplemental Material). The instructions gave no hints about specific effects, their direction, or variation in treatments. In the Discussion, we consider whether the above differences are likely to provide a substantive basis for the discrepancy between our and Carney et al.'s results. Notably, however, as did Carney et al., we found a significant effect of power posing on self-reported feelings of power—which indicates that outcomes on at least one measure were not influenced by the different procedures.

Results

Using two-tailed t tests, we replicated Carney et al.'s finding that participants in the high-power condition self-reported, on average, higher feelings of power than did participants in the low-power condition (mean difference = 0.245, 95% confidence interval, or CI = [0.044, 0.446]), $t(193) = 2.399$, $p = .017$, Cohen's $d = 0.344$.[2] This suggests that the power-posing manipulation worked.

Figure 1 compares mean risk tolerance in the gain domain and mean change in testosterone between the high- and low-power-pose conditions. Two-tailed t tests indicated no significant impact of physical position on either risk-taking in the gain domain (mean difference = −0.033, 95% CI = [−0.085, 0.019]), $t(198) = −1.245$, $p = .215$, Cohen's $d = −0.176$, or changes in testosterone (mean difference = −4.077, 95% CI = [−9.801, 1.647]), $t(196) = −1.405$, $p = .162$, Cohen's $d = −0.200$.[3] We also found no effect for cortisol (mean difference = −0.028, 95% CI = [−0.078, 0.022]), $t(196) = −1.101$, $p = .272$, Cohen's $d = −0.157$. Complementary nonparametric and regression analyses yielded similar results for the whole sample, and for women and men separately (see the Supplemental Material). Neither of the other two behavioral tasks yielded significant differences by condition.

Discussion

Replication is an important tool for identifying the robustness of results, particularly when small sample sizes increase the likelihood of false positives (Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011). Using a much larger sample size but similar procedures as Carney et al. did, we failed to confirm an effect of power posing on testosterone, cortisol, and financial risk taking. We did find that power posing affected self-reported feelings of power; however, this did not yield behavioral effects.

It is possible that subtle differences between the experimental protocols in Carney et al. and those in our study, originally designed as an extension of Carney et al., led to the omission of factors crucial for power poses to influence hormonal levels and behavior. Indeed, a reviewer suggested a few possibilities worth considering. First, prolonged posing time in our study may have caused participants to become uncomfortable, which could have counteracted the effect of power posing. To test this interpretation, we reanalyzed our data using first only those participants who reported the postures to be at least "somewhat comfortable" and then only those who reported the postures to be at least "quite comfortable." This did not substantively change our results (see the Supplemental Material). Moreover, at least one previous study found power-posing effects using similar positions and time intervals as those in our study (Cuddy, Wilmuth, & Carney, 2012; see also Fischer, Fischer, Englich, Aydin, & Frey, 2011; Schubert & Koole, 2009).

Second, certain conditions may need to be present during the power-manipulation phase in order for effects to arise. For example, the reviewer suggested that a social

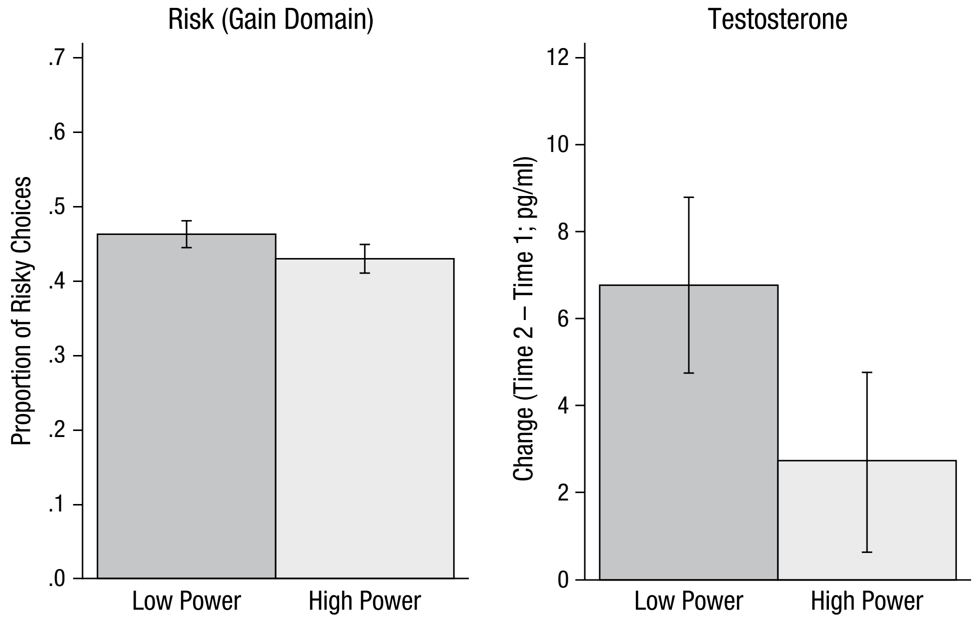

Fig. 1. Mean proportion of risky choices in the gain domain (left) and mean change in testosterone from before the power-pose manipulation (Time 1) to 17 min after the power-pose manipulation (Time 2; right). For each graph, results are shown separately for the high- and low-power-pose conditions. Error bars represent standard errors of the mean.

filler task may facilitate power-posing effects, as suggested by Cesario and McDonald's (2013) Study 1. However, other studies have found effects using tasks without social components (e.g., Fischer et al., 2011, Studies 1 and 3; Yap et al., 2013, Studies 2 and 3). A third possibility involves what subjects were told about the purpose of the study. Most studies that have found effects have employed statements that explicitly obscured the study's purpose with a cover story (Carney et al.; Huang, Galinsky, Gruenfeld, & Guillory, 2011). However, power-posing effects have also been reported in at least one case in which subjects were told that the study was about physical motion and performance (Cuddy et al., 2012), which is similar to the degree of information subjects received in our design.

Hence, while it is certainly plausible that power-posing effects generally arise only under some specific conditions—and our study failed to include one or two of these—it is far from clear from the literature what these conditions are. We conclude that more research is needed to identify the precise conditions necessary for such effects.

Author Contributions

The study was designed by A. Dreber, M. Johannesson, S. Leiberg, E. Ranehill, S. Sul, and R. A. Weber. S. Leiberg and E. Ranehill conducted the study, and E. Ranehill analyzed the data. A. Dreber, M. Johannesson, E. Ranehill, and R. A. Weber wrote the manuscript.

Acknowledgments

We appreciate comments from participants at the 8th Nordic Conference on Behavioral and Experimental Economics and the 4th Annual Xiamen University International Workshop on Experimental Economics. We thank George Loewenstein for helpful comments.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

We gratefully acknowledge financial support from the Jan Wallander and Tom Hedelius Foundation (A. Dreber, M. Johannesson), the Knut and Alice Wallenberg Foundation (A. Dreber), the Swedish Council for Working Life and Social Research (M. Johannesson), and the Swiss National Science Foundation (E. Ranehill, R. A. Weber).

Supplemental Material

Additional supporting information can be found at http://pss.sagepub.com/content/by/supplemental-data

Open Practices

All data and materials have been made publicly available via Open Science Framework and can be accessed at https://osf.io/6akvt/files/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/view/ and http://pss.sagepub.com/content/25/1/3.full.

Notes

1. One participant did not fully comply with the posing instructions. Excluding this participant yielded similar results.
2. We did not have information on feelings of power for 5 participants. Three of these cases were the result of questionnaires not being saved correctly after laboratory maintenance.
3. Hormonal levels were not detectable in one saliva sample. An additional sample was lost during the storage and transportation process.

References

Carney, D. R., Cuddy, A. J. C., & Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological Science*, *21*, 1363–1368. doi:10.1177/0956797610383437

Cesario, J., & McDonald, M. M. (2013). Bodies in context: Power poses as a computation of action possibility. *Social Cognition*, *31*, 260–274.

Cuddy, A. J. C., Wilmuth, C. A., & Carney, D. R. (2012). *The benefit of power posing before a high-stakes social evaluation* (Harvard Business School Working Paper No. 13-027). Retrieved from http://dash.harvard.edu/bitstream/handle/1/9547823/13-027.pdf?sequence=1

Fischer, J., Fischer, P., Englich, B., Aydin, N., & Frey, D. (2011). Empower my decisions: The effects of power gestures on confirmatory information processing. *Journal of Experimental Social Psychology*, *47*, 1146–1154. doi:10.1016/j.jesp.2011.06.008

Huang, L., Galinsky, A. D., Gruenfeld, D. H., & Guillory, L. E. (2011). Powerful postures versus powerful roles: Which is the proximate correlate of thought and behavior? *Psychological Science*, *22*, 95–102. doi:10.1177/0956797610391912

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), Article e124. Retrieved from http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0020124

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, *122*, 1067–1101. doi:10.1162/qjec.122.3.1067

Riskind, J. H., & Gotay, C. C. (1982). Physical posture: Could it have regulatory or feedback effects on motivation and emotion? *Motivation and Emotion*, *6*, 273–298. doi:10.1007/BF00992249

Schubert, T. W., & Koole, S. L. (2009). The embodied self: Making a fist enhances men's power-related self-conceptions. *Journal of Experimental Social Psychology*, *45*, 828–834. doi:10.1016/j.jesp.2009.02.003

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366. doi:10.1177/0956797611417632

Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, *64*, 211–220. doi:10.1037/0022-3514.64.2.211

Yap, A. J., Wazlawek, A. S., Lucas, B. J., Cuddy, A. J. C., & Carney, D. R. (2013). The ergonomics of dishonesty: The effect of incidental posture on stealing, cheating, and traffic violations. *Psychological Science*, *24*, 2281–2289. doi:10.1177/0956797613492425

Default Question Block

Who are you? This is just for organization purposes. Your scores will not be attached to your name in the assessment report.

- ○ Jacob
- ○ Kristin
- ○ Erica

Please select the three-digit identifier on the assignment you are currently evaluating.

[⌄]

Instructions for raters: If participants leave relevant questions blank, please do not give them a rating (i.e., do not give them a "Not at all" - just leave blank)

Below is the rubric for the critical thinking institutional learning outcome (from the university). Please consider the whole assignment when evaluating each student in these areas. That said, specific questions speak to each area better than others.

Explanation of issues: Q1 & Q2
Use of evidence: Q4 & Q6
Context, assumptions: Q7
Alternative viewpoints: Q8
Statement of position: Q7 & Q9
Conclusions, implications, and consequences: Q9

Explanation of issues

4 = Explanation stated clearly and provides all relevant information necessary for full understanding.

3 = Explanation stated less clearly and/or provides mostly relevant information necessary for full understanding.

2 = Explanation stated provides some relevant information necessary for understanding.

1 = Explanation too weak for necessary understanding or not provided.

Use of evidence

4 = Provides sufficient information to support claims and conclusions made.

3 = Provides some information to support claims and conclusions made.

2 = Provides little information to support claims and conclusions made.

◯

1 = Lacks information to support claims and conclusions made.

◯

Context, assumptions

4 =Thoroughly analyzes strengths and weaknesses of one's own and others' assumptions; carefully evaluates influence of context.

◯

3 = Analyzes strengths and weaknesses of one's own and others' assumptions; evaluates context.

◯

2 = Minimally analyzes strengths and weaknesses of one's own and others' assumptions; minimally evaluates context.

◯

1 = Fails to analyze strengths and weaknesses of one's own and others' assumptions; does not evaluate context.

◯

Alternative viewpoints

4 = Carefully evaluates all relevant alternative viewpoints.

○

3 = Evaluates most of the relevant alternative viewpoints.

○

2 = Evaluates some of the relevant alternative viewpoints.

○

1 = Evaluates little/none of the relevant alternative viewpoints.

○

Statement of position

4 = States a clear position that is valid, original and/or innovative, as appropriate.

○

3 = States a relatively clear position that has some validity, originality and/or innovation, as appropriate.

○

2 = States a position that lacks validity, originality, and/or innovation.

○

1 = Does not state a position.

○

Conclusions, implications, and consequences.

4 = Conclusions, implications, and consequences flow from student's analysis.

○

3 = Conclusions, implications, and consequences generally flow from student's analysis.

○

2 = Conclusions, implications, and consequences minimally flow from student's analysis.

○

1 = Conclusions, implications, and consequences do not flow from the student's analysis.

○

Here are a few additional questions.

How well did the student explain the main topic? (Q1)?

Not at all

○

Slightly

○

Somewhat

○

Moderately

○

Extremely

○

How well did the student explain the importance of the main topic (Q1)?

Not at all

○

Slightly

○

Somewhat

○

Moderately

○

Extremely

○

How well did the student explain the importance of replication (Q2; Q9)?

Not at all

○

Slightly

○

Somewhat

Moderately

Extremely

How well did students describe the hypotheses from Carney et al. (2010) (Q3)?

Not at all

Slightly

Somewhat

Moderately

Extremely

How well did the students describe the operationalization of power posing (Q5)?

Not at all

Slightly

Somewhat

Moderately

○

Extremely

○

Did the students accurately summarize the results from Carney et al (2010) (Q6)?

Not at all

○

Slightly

○

Somewhat

○

Moderately

○

Extremely

○

Did the students accurately assess whether Ranehill et al. (2015) replicated Carney et al. (2010) (Q6)?

Not at all

○

Slightly

○

Somewhat

○

Moderately

○

Extremely

○

Did the students accurately assess whether the results supported the authors' hypotheses (Q6)?

Not at all

○

Slightly

○

Somewhat

○

Moderately

○

Extremely

○

How well did the students design a new study to test power posing (Q10)?

Not at all

○

Slightly

○

Somewhat

○

Moderately

○

Extremely

○

To what degree did their study address the potential limitations of Carney et al. (2010) and Ranehill et al. (2015) (Q10)?

Not at all

○

Slightly

○

Somewhat

○

Moderately

○

Extremely

○

Do you suspect AI was used to generate this student's responses?

○ yes
○ no
○ maybe

To what extent do you suspect AI was used to generate this student's responses?

Not at all (e.g., no AI was used to generate responses)

○

Somewhat (e.g., for editing purposes only)

○

Quite a bit (e.g., idea generation, some response writing)

○

Completely (e.g., all responses were AI-generated)

○