Reply

Psychological Science
2018, Vol. 29(4) 656–666

Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797617746749
www.psychologicalscience.org/PS
SAGE

P-Curving a More Comprehensive Body of Research on Postural Feedback Reveals Clear Evidential Value for Power-Posing Effects: Reply to Simmons and Simonsohn (2017)

Amy J. C. Cuddy[1], S. Jack Schultz[2], and Nathan E. Fosse[3]
[1]Department of Psychology, Harvard University; [2]Harvard Business School, Boston, Massachusetts; and [3]The Institute for Quantitative Social Science, Harvard University

Received 8/25/16; Revision accepted 11/10/17

Several *p*-curve analyses based on a systematic review of the current scientific literature on adopting expansive postures reveal strong evidential value for postural-feedback (i.e., *power-posing*) effects and particularly robust evidential value for effects on emotional and affective states (e.g., mood and evaluations, attitudes, and feelings about the self). These findings, based on 55 studies, stand in contrast to those of Simmons and Simonsohn (2017), whose results from a *p*-curve analysis of an older and less comprehensive set of 34 studies led them to conclude that those findings do not possess evidential value.

Background

In a 2015 *Psychological Science* Commentary, Carney, Cuddy, and Yap presented a narrative review of the psychological feedback effects of adopting expansive (vs. neutral or contractive) nonverbal postures (postural feedback). Their Commentary, in which they included "all published tests (to [their] knowledge)" (p. 657), had four primary goals. The first aim was to provide a theoretical and methodological summary of available experimental studies of the effects of postural feedback on various psychological outcomes, including cognitive, emotional, behavioral, and physiological measures. The second was to compare and contrast similarities and differences across these studies and a conceptual replication attempt of the main study in Carney, Cuddy, and Yap (2010), which was reported in a 2015 *Psychological Science* Commentary (Ranehill et al., 2015). Third, they wanted to identify possible moderators of postural-feedback effects, and finally, they aimed to determine promising avenues for future research. It was neither intended to be nor was presented as a systematic review and meta-analysis of the literature.

In a 2017 *Psychological Science* Commentary and in a 2015 Data Colada blog post, Simmons and Simonsohn submitted the studies listed in the Carney, Cuddy, and Yap (2015) narrative review to a *p*-curve analysis, a meta-analytic technique described below, and concluded that their results do not support the existence of a real effect of power posing and that "the existing evidence is too weak to justify a search for moderators or to advocate for people to engage in power posing to better their lives" (Simmons & Simonsohn, 2017, pp. 690–691).

Overview

We conducted a series of *p*-curve analyses following Simonsohn, Nelson, and Simmons's (2015) rules of *p*-curving and using a systematically selected, comprehensive, and updated set of published studies of power posing. These analyses yielded starkly different results from those of Simmons and Simonsohn (2015, 2017): (a) evidential value for postural feedback across aggregated effects; (b) evidential value for a clearly specified, theoretically important, single effect—feelings of power—which was omitted from the *p*-curve figures presented by Simmons and Simonsohn; and (c) remarkably strong

Corresponding Author:
Amy J. C. Cuddy, Harvard University, Department of Psychology, William James Hall, 33 Kirkland St., Cambridge, MA 02138
E-mail: amycuddy@fas.harvard.edu

evidential value for a well-defined, theoretically important category of effects from the same set of studies identified in our systematic review—all measures of feelings, including emotions, affect, mood, and evaluations, attitudes, and feelings about the self.

In this Reply, we discuss our new analyses and how it is that two groups of researchers, strictly applying the same analytic technique, can reach sharply conflicting conclusions about the extent to which an area of research does or does not contain evidential value.

P-Curving

First, we present a brief primer on *p*-curving, a technique that was introduced by Simonsohn, Nelson, and Simmons in 2014: A *p*-curve is the distribution of statistically significant ($p < .05$) *p* values selected for each study (one effect per study for the *main curve* and one effect per study for the *robustness curve*) in a given set of studies that is defined by the "*p*-curver." On the basis of the distribution of *p* values, the authors argue, one can "distinguish between sets of significant findings that are likely versus unlikely to be the result of selective reporting" (Simonsohn et al., 2014, p. 535), determining whether the body of research possesses evidential value (a right-skewed curve), inadequate evidential value (a flatter distribution than we would expect to find if the underlying studies had an average power of 33%, a threshold that Simmons and Simonsohn have described as arbitrary but justifiable), or indications that *p*-hacking was used to achieve statistical significance (a left-skewed curve).[1] According to Simonsohn et al. (2014), *p*-curving also produces an estimate of the average statistical power of the studies that corrects for selective reporting. Recently, Simonsohn, Simmons, and Nelson (2015) revised their methods, including not only tests of skew for all *p*s < .05, but also for *p*s < .025.

All conclusions drawn from a *p*-curve analysis are necessarily constrained by the content of the input. In their 2017 Commentary and 2015 blog post, Simmons and Simonsohn restricted their analysis to the studies cited by Carney et al. (2015). Thus, their *p*-curve combined widely disparate dependent variables (DVs; e.g., pain tolerance, thought abstraction, self-reported vengeful intention, gambling, hormonal changes, eating behavior). Moreover, the Simmons and Simonsohn *p*-curve omitted many relevant studies. As already noted, they also excluded self-reported feelings of power as a DV. We report *p*-curve analyses in which we overcame those limitations by systematically updating the set of studies and further examining clearly defined and theoretically meaningful subsets of effects. We believe that for *p*-curving to produce the most accurate and useful findings and conclusions, it must be applied using the best available evidence to test clearly specified a priori research hypotheses regarding well-defined effects, the aim of our analyses.

The Present Analyses

In the present analyses of the postural-feedback literature, we aim to answer three meta-analytic questions that we defined a priori. First, does a systematic review of the literature pertaining to studies of the feedback effects of adopting expansive versus contractive (or neutral) postural manipulations, consistent with standards established by *p*-curves, possess evidential value? Second, does the effect of postural feedback on a clearly specified, theoretically important single outcome, feelings of power, possess evidential value? Third, does the effect of postural feedback on a well-defined, theoretically meaningful and coherent category of findings from the main curve, those measuring other emotional and affective states (e.g., emotions, affect, mood and evaluations, attitudes, and feelings about the self), possess evidential value?

Systematic review and aggregate analysis

We began by conducting a systematic review of the literature with the aim of identifying the complete set of published empirical studies of power posing up to December 20, 2016. While narrative reviews provide a qualitative description of a body of literature (e.g., Carney et al., 2015), systematic reviews are based on a priori research questions regarding the evaluation of a body of theoretically relevant literature, which then guide careful and comprehensive study inclusion and exclusion (e.g., see Cooper, 2016; Uman, 2011). In our first *p*-curve analysis, our goal was to generate a *p*-curve based on our comprehensive search of the power-posing literature and to compare it with the Simmons and Simonsohn curve. Any differences would provide information about whether and how the content of the included studies and effects can affect the results and conclusions. This analysis addressed our first methodological question: How do sample-selection decisions influence the *p*-curve results and conclusions regarding this broad set of findings?

Feelings-of-power analysis

For our next analyses, acknowledging that limited conclusions can be drawn from these aggregate tests of a heterogeneous set of effects, we refined the inputs to address our second methodological question: How does the undifferentiated aggregation of widely disparate effects into a single *p*-curve influence the conclusions that can be drawn about this broad set of findings?

Thus, in our second analysis, we *p*-curved one causal association between expansive posture and a clearly defined, theoretically meaningful single measure: the effect of postural expansiveness on feelings of power. As theorized by Carney et al. (2010, 2015) and by scores of social psychologists who study power (see Galinsky, Rucker, & Magee, 2015, for a review), feeling powerful is an intrinsically consequential, theoretically important, fundamental outcome. Feelings are core to the field's most popular self-definition: "Social psychology is the scientific study of how people's thoughts, feelings, and behaviors are influenced by the actual, imagined, or implied presence of others" (Allport, 1985, p. 3). As Wegner and Gilbert (2000) explained,

> the center around which modern social psychology actually turns is *the understanding of subjective experience* . . . social psychology is intimately concerned with the scientific understanding of what it is like to be a person—why our existence at this moment and in time and space *feels* the way it does. (p. 1)

And it is a fait accompli that emotions and affect influence cognitive, behavioral, physiological, and other outcomes; this is, in fact, one of the key principles underlying much of social psychology, in particular, and the social sciences in general (e.g., Bertrand & Mullainathan, 2001; Russell, 2003; Wegner & Gilbert, 2000). For example, subjective states and experiences, such as feelings of agency, happiness, and evaluations of the self, predict objective measures of behavior, health, and general well-being (e.g., Aneshensel, Phelan, & Bierman, 2013; Luhmann, Hoffman, Eid, & Lucas, 2012). And emotion theorists have long demonstrated the primacy of affect as preceding and motivating both cognition and behavior (see Zajonc, 1998, for a review).

Specific to the psychology of power, hundreds of studies by researchers, including Fiske, Galinsky, Guinote, Inesi, Keltner, Magee, Overbeck, P. Smith, have firmly established that subjective feelings of power influence both cognition, behavior, and physiological outcomes, including stereotypes, resistance to influence, creativity and authenticity, physical and mental performance, self-regulation, goal pursuit, physiology, health, general well-being, and many others (e.g., Galinsky et al., 2015).

Moreover, theories of body-mind feedback investigating various effects of nonverbal behavior on people's emotional and affective states date back to William James's (1896/1994) late-18th-century theories of emotion and ideomotor action (see also Laird & Lacasse, 2014). Evidence that adopting postural expressions of emotions not only reflects but also shapes emotions contributes to a foundational area of social psychological theory (Niedenthal, 2007). However, Simmons and Simonsohn excluded tests of effects of postural manipulations on self-reported feelings of power from their *p*-curve analysis, arguing that such measures are merely manipulation checks. We disagree. In the seminal Carney et al. (2010) article, self-reported power was repeatedly described as a DV of primary interest, from the abstract through to the discussion. The systematic review reported in the current article yielded 14 studies that measured feelings of power; 12 of those studies treated feelings of power as measures of theoretical interest. Only 2 studies (Cuddy, Wilmuth, Yap, & Carney, 2015, and Ranehill et al., 2015), characterized self-reported feelings of power as a manipulation check, but we believe that was in error. Certainly, the vast majority of studies in this literature have not described feelings of power as a mere manipulation check. In fact, in some studies (e.g., Park, Streamer, Huang, & Galinsky, 2013, Studies 2a and 2b), feelings of power was the only DV and was explicitly presented as the key outcome, not as a manipulation check (Simmons and Simonsohn excluded those studies from their main curve). So, it is not normative in this literature to treat feelings of power as a manipulation check. It would simply not make sense to exclude feelings of power from a *p*-curve analysis of this literature.

EASE analysis

Our third and fourth *p*-curve analyses examined what happens when we reach beyond feelings of power to look at the evidential value of postural-feedback effects on other feelings—emotions, affect, mood, and evaluations, attitudes, and feelings about the self. Focusing on emotion and affect makes sense for several reasons. First, much of the research on postural feedback, which is theoretically grounded in the relationship between nonverbal expressions and emotion, has naturally focused on the effects of expansive postures on emotional and affective states (as opposed to cognition and behavior); it is of primary theoretical interest. Second, it allows us to address concerns about undifferentiated aggregation without limiting the analysis to a single emotion, feelings of power. Third, by including the entire set of emotion- and affect-related outcomes while excluding feelings of power, we can confront questions about whether postural-feedback effects are merely demand effects. The remaining set of emotion- and affect-related outcomes includes findings obtained in procedures that seem unlikely to be susceptible to demand characteristics. In some such studies, there were no obvious cues as to what sort of response was "demanded" (e.g., mood recovery, changes in various discrete emotion states, changes in negative affect, assignment of valence to a series of thoughts following an open-ended thoughts-listing task) or the outcome variables seemed difficult to control or to fake (e.g., speed of retrieval of positive and negative personal memories, mood recovery, ability to recall positive vs. negative words from a list presented earlier in the study, changes to discrete emotion states embedded in a long list of emotions).

Applying a systematic coding procedure, we limited inclusion in the third curve to all emotion- and affect-related outcomes, which we refer to as emotion, affect, and self-evaluation (EASE) variables, while excluding feelings of power. Excluding feelings of power from our assessment of evidential value for postural feedback on emotion-related effects makes that analysis more conservative.

While EASE variables represent a theoretically meaningful subset of the effects included in our first analysis of aggregated outcomes, non-EASE measures do not; they are the theoretically heterogeneous effects that remain after extracting the theoretically coherent set of EASE variables. Additionally, conclusions based on *p*-curve analyses of non-EASE variables require the same caution as for the theoretically heterogeneous set of measures involved in the *p*-curve analysis by Simmons and Simonsohn (2017). We describe our categorization methods for EASE and non-EASE variables in further detail below.

Analytic approach

Taking a conservative analytic approach, we *p*-curved the postural-feedback literature, applying the selection criteria and statistical methods prescribed by Simmons and Simonsohn with only two distinct differences in analytic approach. First, we defined our questions a priori and systematically gathered all available data relevant to the questions at hand. Second, as described above, we conducted several *p*-curve analyses—one for the aggregated outcomes, as Simmons and Simonsohn did, one for feelings of power, and a pair for EASE and non-EASE variables, respectively. We present all *p*-curve analysis results below.

In addition to the original 20 articles (34 studies) assessed by Carney et al., 2015, we performed a systematic literature search for additional studies. We searched for peer-reviewed studies using the Harvard Hollis+ platform. From the known literature, we identified the following keywords (e.g., "power," "dominance," "pride," "shame," "expansive," "open," "upright," "contractive," "slouched," "hunched," "closed," "pose," "posture"), which yielded numerous results.

All studies had to feature a postural manipulation that (directly or indirectly) induced expansive or contractive nonverbal postures, consistent with prior definitions (Carney et al., 2010, 2015). We included only studies with postural manipulations that involved a modification of the orientation and openness of the chest (or torso), shoulders, or both. Studies that manipulated only head orientation (e.g., chin and head down vs. chin and head upright), for example, were not included. Postural manipulations could also include changes to the orientation of the arms, legs, head, and neck, though each of these elements was not considered sufficient on its own. Because postural expansiveness is a continuous spectrum, one of the posture conditions had to be more expansive relative to the others. Studies that featured sitting or standing postural effects were included, while studies testing the effects of supine postures or movement (i.e., walking or dancing) were excluded. (For additional information about the methodology of our systematic literature review, see the materials on the Open Science Framework [OSF].)

The literature search produced an additional 21 studies that met all criteria for inclusion. Those 21 studies, added to the original 34 studies, resulted in a sample of 55 studies. All 55 studies met our inclusion criteria, and all were accounted for in the results of our systematic literature search. From each study, we selected the appropriate statistic (or statistics) on the basis of the rules provided by the *p*-curve guide and the selection criteria used by Simmons and Simonsohn. We selected the first reported hypothesis pertaining to postural-feedback effects when the hypothesis was clearly stated. For our robustness curve, we followed the practices used by Simmons and Simonsohn in (a) carrying over main results where the *p*-curve guide did not require a specific alternative statistic, (b) selecting the specified alternative statistic in cases where the *p*-curve guide required it, and (c) including the appropriate statistic for a second hypothesized effect in cases where there were multiple hypothesized effects. When the article presented multiple hypothesized effects pertaining to postural feedback, we included the second statistic that was explicitly hypothesized or the second reported statistic pertaining to a general hypothesis. (See the disclosure table in our supplemental materials at the OSF for detailed information regarding all statistics that were selected and included in each *p*-curve.)

The EASE *p*-curve required reliable categorization of variables as EASE or non-EASE. Five experts coded the variables: the first two authors of this article and, to ensure objectivity, three additional expert coders, all of whom are social psychologists and tenured professors at research universities but none of whom does research on postural feedback. Coders were contacted by an e-mail that included a link to an online survey. They were asked to categorize all measures that were drawn from our systematically selected set of peer-reviewed experiments testing postural-feedback effects and that were included in the aggregate curve. Coders were provided with the list of dependent measures, named as they were by the original researchers, along with excerpts from the original articles that described exactly how the variables were operationalized. Including both the names and operationalizations of the variables ensured that the coding was indeed based on what was actually measured, given that there are sometimes discrepancies between the conceptual variable and the operationalized variable. Coders were asked to identify "measures of emotions, affect, mood, and evaluations, attitudes, and feelings about the self (i.e., self-evaluations)" and instructed that "measures can be explicit or implicit, direct or indirect, but they should be primary measures, as opposed to correlates, of an EASE construct." The intraclass correlation coefficient (ICC), which we calculated to assess interrater agreement, indicated excellent reliability,[2] average ICC = .92 (95% confidence interval = [.88, .95]). The final categorizations of variables as EASE or non-EASE were determined by majority rule. The final EASE set included such variables as retrieval of positive and negative memories, mood recovery, changes in specific emotion states, recall of positive and negative words from lists presented earlier in the study, and self-evaluations, demonstrating that the effects of postural feedback on affective variables clearly extend beyond feeling powerful. (For more details about the coding of EASE and non-EASE variables, refer to the disclosure table in our supplemental materials at the OSF.)

Results

Evidential value of postural-feedback effects on aggregated variables

Our first *p*-curve analysis, based on a systematic literature review that aimed to include all published empirical tests of power-posing manipulations as of December 20, 2016, comprises 53 statistical results and clearly demonstrates that the postural-feedback literature contains strong evidential value (Fig. 1a). This *p*-curve serves as the comparator to the main curve presented by Simmons and Simonsohn. In fact, we found evidential value both in our main and robustness *p*-curves, as well as with the half *p*-curve featured in the latest version of the *p*-curve app (Version 4.05), which assesses evidential value among studies with *p*s below the median.[3]

A literature is determined to contain "evidential value if either the half *p*-curve . . . is significantly right-skewed at the 5% level, or if both the half and full *p*-curve are significantly right-skewed at the 10% level" (Simmons & Simonsohn, 2017, p. 690). In this case, all conditions for evidential value were met, showing clear right skew for both p_{half} (< .0001) and p_{full} (< .0001) in our main *p*-curve and for both p_{half} (< .0001) and p_{full} (< .0001) in our robustness *p*-curve. Second, the observed *p*-curve was compared with "what would be expected when studies have an average power of only 33%" (Simmons & Simonsohn, 2017, p. 690); a $p < .05$ for the full *p*-curve (or $p < .10$ for the full and binomial *p*-curves) would indicate a flatter curve than we would expect when the included studies have an average power of 33% and an absence of evidential value. The results of this analysis did not meet any of the criteria for an absence of evidential value, $p_{\text{full}} = .8003$, $p_{\text{binomial}} = .0805$ for our main *p*-curve; $p_{\text{full}} = .9036$, $p_{\text{binomial}} = .1184$ for our robustness curve. Third, we found that the estimate of average power for the set of studies was 44% in our main curve and 49% in our robustness curve (compared with the Simmons and Simonsohn estimates of 5% in both curves). When submitting this systematically identified current set of studies to *p*-curve analysis, we found that the literature on postural feedback possesses evidential value.

In sum, the results of our aggregate *p*-curve, which included a systematically identified and comprehensive set of studies, demonstrate that this literature possesses evidential value. This finding sharply diverges from the results and conclusions of Simmons and Simonsohn's *p*-curve analysis, which failed to show evidential value. (Although note that no *p*-curve analysis by either set of authors yielded results that were left skewed or that suggested that the existing evidence was *p*-hacked.)

Evidential value for postural-feedback effects on feelings of power

In Figure 1b, we present the *p*-curve analysis for postural-feedback effects on feelings of power, which clearly demonstrates that these effects possess evidential value. First, the analysis yielded strong evidence of right skew for both p_{half} (.0009) and p_{full} (< .0003).[4] Second, the results of the tests for flatness did not meet any of the criteria for an absence of evidential value, $p_{\text{full}} = .8803$, $p_{\text{binomial}} = .6652$. Third, the estimated average power for the specific feelings of power outcome was 58%, higher than the estimated average power for our aggregate curve. This *p*-curve analysis shows strong

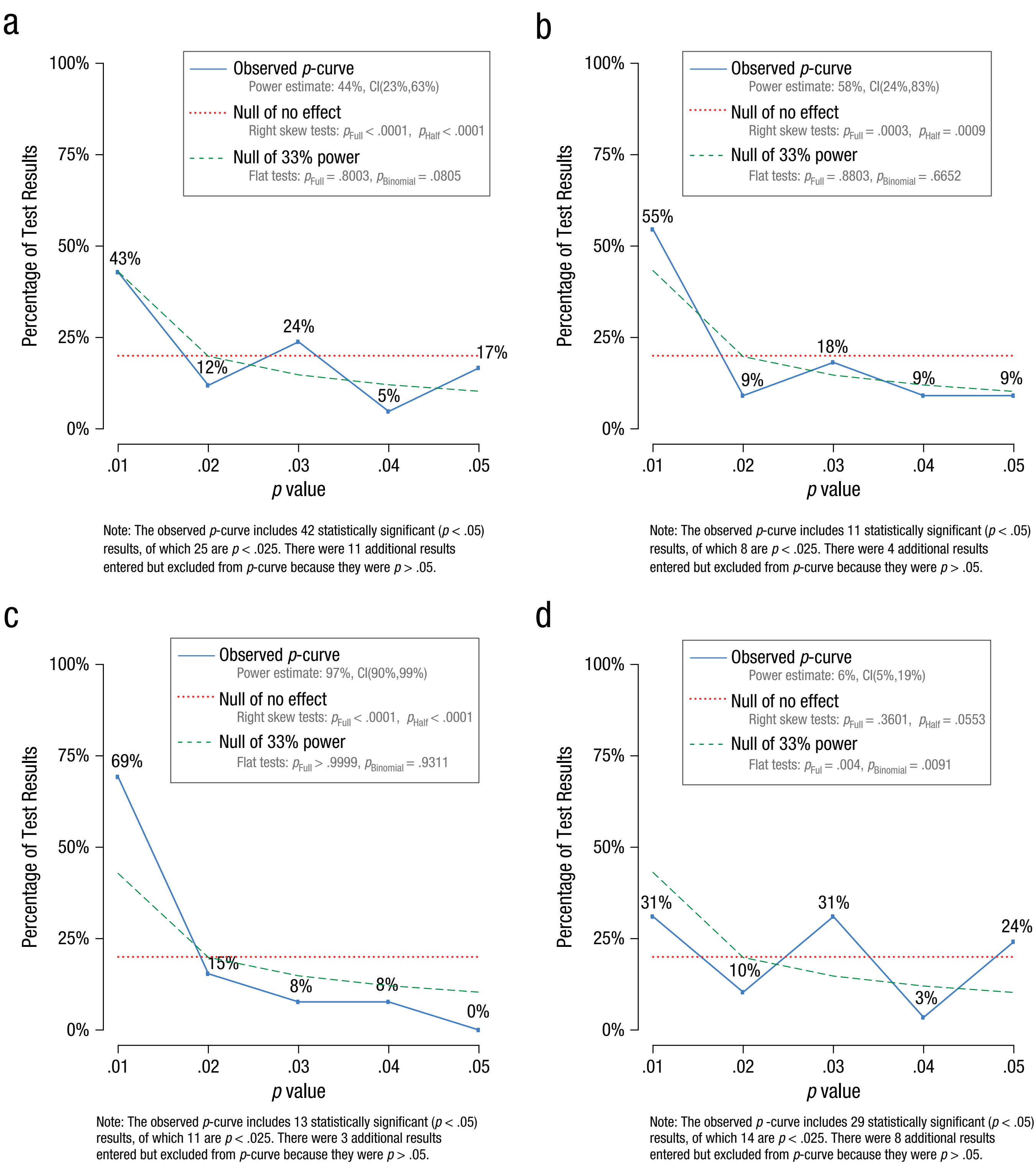

Fig. 1. *P*-curves of data from the postural-feedback literature: (a) the aggregate *p*-curve for postural-feedback effects; (b) the *p*-curve for feelings of power; (c) the *p*-curve for emotion, affect, and self-evaluation (EASE) variables; and (d) the *p*-curve for non-EASE variables. The graphs were generated using the *p*-curve app (Version 4.05; Simmons & Simonsohn, 2017).

evidential value for postural-feedback effects on feelings of power—a clearly specified and theoretically important single outcome.

Evidential value for postural-feedback effects on EASE and non-EASE variables

EASE variables. In Figure 1c, we present the *p*-curve analysis for postural-feedback effects on EASE variables, which clearly reveals robust evidential value. First, the analysis yielded very strong evidence of right skew, $p_{\text{half}} < .0001$, $p_{\text{full}} < .0001$. Second, the results of the tests for flatness did not meet any of the criteria for an absence of evidential value, $p_{\text{full}} > .9999$, $p_{\text{binomial}} = .9311$. Third, the estimated average power for the EASE variables was extremely high (97%), well exceeding the estimated average power of both the aggregate and feelings-of-power curves.

This *p*-curve analysis of a well-defined, theoretically important category of postural-feedback effects—EASE variables—demonstrates very strong evidential value. Expansive versus contractive posture affects not only how powerful people feel but how people feel on a wide variety of other emotion- and affect-related outcomes.

Non-EASE variables. In Figure 1d, we present the *p*-curve analysis for postural-feedback effects on non-EASE variables, a miscellaneous subset of the statistics featured in the main curve (Fig. 1a) of our aggregate test. These are the theoretically heterogeneous "leftovers" after extracting the EASE effects. The test for right skew was marginally significant for the half curve ($p_{\text{half}} = .0553$) and nonsignificant for the full curve ($p_{\text{full}} = .3601$). Additionally, tests for a null of 33% power indicate an absence of evidential value ($p_{\text{full}} = .0040$, $p_{\text{binomial}} = .0091$).

Although a significant half *p*-curve would be adequate to determine that a set of studies possesses evidential value (Simonsohn, Simmons, & Nelson, 2015), the flatness tests clearly failed to reject the null of 33% power. Thus, for this nebulous set of non-EASE variables, the *p*-curve analysis yielded very weak support for the existence of evidential value.

Discussion

These analyses bring to light several critical discoveries about the existing postural-feedback literature. When we included a comprehensive and current set of evidence, comprising 55 studies identified through a systematic review, *p*-curve analyses revealed (a) clear evidential value for postural feedback on an aggregated set of effects; (b) strong evidential value for a clearly specified and theoretically important single outcome, feelings of power; (c) very strong evidential value for a well-defined and theoretically important category of other feelings effects (i.e., EASE variables, which did not include feelings of power); and (d) an absence of evidential value for the theoretically heterogeneous non-EASE effects that remained after separating out the EASE variables. Our findings also suggest that *p*-curving is likely to yield more accurate and informative results when researchers address the following practices: (a) faulty sample-selection decisions and (b) undifferentiated aggregation of disparate effects. When these practices are not adequately addressed, *p*-curve conclusions can lead to misguided dismissals of broad areas of research.

Strong evidential value for postural-feedback effects, particularly for emotions

Our *p*-curve analyses of emotion- and affect-related outcomes yielded robust evidence that postural feedback influences self-reported affective states. First, we found strong evidential value for a precisely specified outcome, feelings of power. That finding converges with a recent Bayesian meta-analysis of a new set of studies that, as described by Cesario, Jonas, and Carney (2017), "showed a reliable non-zero effect on felt power" (p. 2).[5] Presenting their results, Gronau et al. (2017) write that "our meta-analysis yields very strong evidence for an effect of power posing on felt power" (p. 123). In the set of studies presented in our analyses, 11 demonstrated a significant effect of power posing on feelings of power; that does not include studies from 2017, which would increase the total number of replications. Together, the collective evidence provides strong support for the effect of postural feedback on feelings of power. From our theoretical perspective, an expansive posture is a universal expression of power, and adopting such a pose leads people to feel more powerful. The finding of evidential value for self-reported feelings of power directly supports that claim. Moreover, we believe that even transient feelings of power can have long-lasting consequences for people's lives (e.g., Galinsky et al., 2015).

The robust evidential value for postural-feedback effects on EASE variables—emotions, affect, and self-evaluations—is particularly illuminating. These findings from the present set of studies provide convincing evidence that postural manipulations affected subjects' specific emotions, affect, mood recovery, retrieval and recall of positive versus negative memories, and self-evaluations, demonstrating that the effects of postural feedback on affective variables clearly extend beyond causing people to feel more powerful. It is worth noting that the direction of most of the EASE effects is consistent with Keltner, Gruenfeld, and Anderson's (2003) approach-inhibition theory of power: Power activates the behavioral approach system (e.g., recall of more positive than negative words from memory, improved general mood and mood recovery, increased feelings of strength, decreased feelings of fear).

Many studies that are featured in our EASE curve were likely robust to potential demand characteristics, since they used a single- or double-blind study design, deception, or "non-deceptive obfuscation" (Zizzo, 2010, p. 75); tested hypotheses that were simply not intuitive to participants (e.g., mood recovery, changes in various discrete emotion states, changes in negative affect, assignment of valence to a series of thoughts following an open-ended thoughts-listing task); or directly tracked the extent to which participants guessed the hypothesis in exit interviews (which showed that they did not). Some studies were more resilient to demand effects because responses were implicit or otherwise difficult for participants to control (e.g., speed of retrieval of positive and negative personal memories, mood recovery, ability to recall positive and negative words from a list presented earlier in the study, changes to discrete emotion states embedded in a long list of emotions), responses were embedded in a broader survey instrument (e.g., changes in discrete self-reported emotions embedded in a long list of emotions), or, as demonstrated in recent research on demand effects in survey research, participants likely varied in their orientation such that some would have wanted to confirm the hypothesis and some to disconfirm it, and others would have been indifferent (Mummolo & Peterson, 2017). Citations for each of these examples are listed in our supplemental materials at the OSF. Our assessment of the input for our EASE *p*-curve analysis, the strongest *p*-curve presented, is that it is unlikely that these postural-feedback effects are demand effects, given the study designs and the latest research on demand characteristics.

In contrast to the EASE *p*-curve, the non-EASE *p*-curve comprises a theoretically heterogeneous, noncohesive collection of effects (e.g., number of calories consumed at a meal, pain threshold, vengeful intentions, performance on creativity tasks, hormones, beliefs about religion, performance in a job interview, gambling), making any results, whether they indicate a presence or absence of evidential value, difficult to interpret. Removing the EASE variables flattens the curve for the remaining effects, which could indicate that evidential value for behaviors and hormones is weak. This interpretation is consistent with the mostly null results of the set of studies in a recent special issue of *Comprehensive Results in Social Psychology* (Jonas et al., 2017) that measured effects of power posing on various behavioral outcomes. However, many of the non-EASE effects include nonbehavioral or hormonal effects, such as cognitive abilities, creativity, and attitudes; the evidence for these effects seems to be stronger. It is also worth noting that the non-EASE effects include measures that are susceptible to demand characteristics, such as gambling, pain tolerance, and action tendencies in hypothetical scenarios. There is also a need for experimental tests of incremental or longitudinal effects of adopting expansive postures over time on various outcomes. Right now, we are not aware of any such research. As more studies are conducted and published, it will become easier for researchers to analyze other theoretically meaningful subsets of effects, such as hormonal effects, performance under stress, risk preferences, and cognitive abilities. Such analyses of these subsets will continue to enhance the definition of this picture.

What do these analyses tell us about the evidence for postural-feedback effects? Given the present *p*-curve analyses, as strictly interpreted in accordance with the rules of *p*-curving, one must first conclude that the current literature on postural feedback does possess evidential value. By systematically identifying and analyzing meaningful subsets of effects, *p*-curving begins to give more definition to our findings and to the overall picture: The existing effects of postural feedback on feelings possess extremely strong evidential value. As the overall body of studies grows, it will become easier to analyze other meaningful subsets, such as cognitive measures, performance behaviors, and psychophysiological outcomes. Combining these more-focused meta-analyses of meaningful categories of effects with new, theory-driven studies that employ improved methods (e.g., preregistration of a priori hypotheses, larger samples, more precise hormone-measurement instruments and methods) and that come from various disciplines will advance and refine our theoretical understanding of postural feedback—and the same will be true for other areas of research. This will lead to the identification of contextual variables that moderate effects and help us to resolve conflicting evidence from studies of some of the specific effects, such as hormones and risk taking, which have produced both significant and null effects. The analyses do not tell us, however, about the extent to which there is evidential value for other meaningful categories of effects, which individual postural-feedback effects are most robust, which of them might be false positives, and how these complex relationships among posture and these many different variables may be affected by various moderators. It should go without saying that these curves are not the final curves. No meta-analysis can be the final meta-analysis, because results hinge entirely on the content of what is included, and that content will continue to grow and change. Science is cumulative by nature.

The roles of sample-selection decisions and undifferentiated aggregation of effects

How did two groups of researchers reach such discrepant findings and conclusions about the same area of research? Our analyses reveal two of the practices that contributed. First, we addressed the issue of sample-selection decisions that may lead to an incomplete or non-representative set of studies and effects for inclusion in the analysis. Differences between Simmons and Simonsohn's selections and our selections gravely influenced the results of the Simmons and Simonsohn analysis and the conclusions they drew from those results, which dramatically differed from the results and conclusions from our analyses, which were guided by our a priori systematic review of all available literature. As Simonsohn et al. (2014) wrote in their seminal article, "for inferences from *p*-curve to be valid, studies and *p*-values must be appropriately selected" (p. 535). Note that sample selection is not limited to the selection of studies; it can also extend to the selection of effects from each study, particularly when a study includes multiple DVs that are equally weighted by the primary researchers (e.g., from Carney et al., 2010, p. 1364: "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"). Second, our *p*-curve analyses of feelings of power and EASE variables underscore our concerns that undifferentiated aggregation can muddy the waters, making it difficult to draw accurate conclusions from *p*-curve analyses of widely disparate effects. In the present case, the results from Simmons and Simonsohn (2017) mask markedly strong effects of postural expansiveness on feelings of power and on other emotional and affective states.

We are not arguing that the statistical results of Simmons and Simonsohn's *p*-curve analysis (2015, 2017) are incorrect; we are arguing that their results and conclusions, as a result of the practices described above, are misleading with regard to assessments of the evidential value of this area of research. The present *p*-curve results annul Simmons and Simonsohn's (2017) conclusion that "the existing evidence is too weak to justify a search for moderators or to advocate for people to engage in power posing to better their lives" (pp. 690–691). Our findings, including modest support for the general literature on postural feedback and particularly strong support for effects on emotional and affective states, should encourage researchers who are investigating this area to continue doing so.

Action Editor

D. Stephen Lindsay served as action editor for this article.

Author Contributions

A. J. C. Cuddy led the development of this article's conceptual framework. All authors contributed to the analytic approach. Sample selection was led by S. J. Schultz. Statistical analyses were led by N. E. Fosse. A. J. C. Cuddy and N. E. Fosse wrote the first draft of the manuscript. A. J. C. Cuddy and S. J. Schultz revised the manuscript, with input from N. E. Fosse. All authors approved the final version of the manuscript for submission.

Acknowledgments

We thank the following people for providing constructive, expert input at various times and in various ways to the development of this article: Xiang Ao, Don Carlston, Paul Coster, Naomi Ellemers, Eli Finkel, Susan Fiske, Dan Gilbert, Gary King, Andrew Marder, Allen McConnell, and Terri Vescio. A. J. C. Cuddy and S. J. Schultz were at Harvard Business School when the first few versions of this manuscript were written and submitted; A. J. C. Cuddy is now affiliated with the Department of Psychology at Harvard University, and S. J. Schultz is no longer at Harvard University. N. E. Fosse was a postdoctoral associate at Harvard University's Institute for Quantitative Social Science when the first few versions of this manuscript were written and submitted; he is now at the Harvard University Division of Continuing Education.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/pfh6r/. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617746749. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

Notes

1. We also applied an earlier version of the *p*-curve app that included output on the likelihood of *p*-hacking as indicated by left skew. There was no evidence of *p*-hacking in any analysis of our or Simmons and Simonsohn's *p*-curves. We downloaded and ran a copy of the *p*-curve app R script (Version 4.05, written

by Uri Simonsohn) from www.pcurve.com. We provide a copy of the R script for Version 4.05 along with our disclosure table in our supplemental materials at the Open Science Framework (OSF; https://osf.io/pfh6r/). Earlier versions of the *p*-curve app R scripts are presently unavailable, though *p*-curve app updates are listed at http://www.p-curve.com/app4/versions.php.
2. For ICC guidelines, see Koo and Li (2016).
3. The lowest *p* value featured in our main curve corresponds to a measure of self-reported strength, drawn from the study by Peper, Booiman, Lin, and Harvey (2016), in which participants resisted downward pressure applied to their arms while maintaining an erect or collapsed posture and reported how strong they felt. The study protocol was designed to minimize demand characteristics, and the results show a clear effect on how posture affects felt strength.
4. Given the *p*-curving practice of carrying over the main results where the *p*-curve guide does not require a specific alternative statistic, the feelings-of-power robustness curve does not have any alternative statistics to draw from, so a robustness curve would be entirely redundant with the main curve.
5. The end date of our systematic review was December 20, 2016, and we judged it inappropriate to add studies that we learned of incidentally thereafter because doing so would have undermined the objectivity and integrity of the systematic review (although we provide references to several 2017 studies in our supplemental materials at the OSF).

References

Allport, A. (1985). The historical background of social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 1, pp. 1–46). New York, NY: Random House.

Aneshensel, C. S., Phelan, J. C., & Bierman, A. (2013). The sociology of mental health: Surveying the field. In C. S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health* (pp. 1–19). Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-4276-5_1

Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *The American Economic Review*, *91*, 67–72. doi:10.1257/aer.91.2.67

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

Carney, D. R., Cuddy, A. J. C., & Yap, A. J. (2015). Review and summary of research on the embodied effects of expansive (vs. contractive) nonverbal displays. *Psychological Science*, *26*, 657–663. doi:10.1177/0956797614566855

Cesario, J., Jonas, K. J., & Carney, D. R. (2017). CRSP special issue on power poses: What was the point and what did we learn? *Comprehensive Results in Social Psychology*, *2*, 1–5. doi:10.1080/23743603.2017.1309876

Cooper, H. (2016). *Research synthesis and meta-analysis: A step-by-step approach*. Thousand Oaks, CA: SAGE.

Cuddy, A. J. C., Wilmuth, C. A., Yap, A. J., & Carney, D. R. (2015). Preparatory power posing affects nonverbal presence and job interview performance. *Journal of Applied Psychology*, *100*, 1286–1295. doi:10.1037/a0038543

Galinsky, A. D., Rucker, D. D., & Magee, J. C. (2015). Power: Past findings, present considerations, and future directions. In M. Mikulincer & P. R. Shaver (Eds.), *APA handbook of personality and social psychology* (Vol. 3: Interpersonal Relationships, pp. 421–460). Washington, DC: American Psychological Association. doi:10.1037/14344-016

Gronau, Q. F., Van Erp, S., Heck, D. W., Cesario, J., Jonas, K. J., & Wagenmakers, E.-J. (2017). A Bayesian model-averaged meta-analysis of the power pose effect with informed and default priors: The case of felt power. *Comprehensive Results in Social Psychology*, *2*, 123–138. doi:10.1080/23743603.2017.1326760

James, W. (1994). The physical basis of emotion. *Psychological Review*, *101*, 205–210. (Original work published 1896) doi:10.1037/0033-295X.101.2.205

Jonas, K., Cesario, J., Alger, M., Bailey, A. H., Bombari, D., Carney, D., . . . Tybur, J. M. (2017). Power poses – where do we stand? *Comprehensive Results in Social Psychology*, *2*, 139–141. doi:10.1080/23743603.2017.1342447

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*, 265–284. doi:10.1037/0033-295X.110.2.265

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155–163. doi:10.1016/j.jcm.2016.02.012

Laird, J. D., & Lacasse, K. (2014). Bodily influences on emotional feelings: Accumulating evidence and extensions of William James's theory of emotion. *Emotion Review*, *6*, 27–34. doi:10.1177/1754073913494899

Luhmann, M., Hofmann, W., Eid, M., & Lucas, R. E. (2012). Subjective well-being and adaptation to life events: A meta-analysis. *Journal of Personality and Social Psychology*, *102*, 592–615. doi:10.1037/a0025948

Mummolo, J., & Peterson, E. (2017). *Demand effects in survey experiments: An empirical assessment*. doi:10.2139/ssrn.2956147

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*, 1002–1005.

Park, L. E., Streamer, L., Huang, L., & Galinsky, A. D. (2013). Stand tall, but don't put your feet up: Universal and culturally-specific effects of expansive postures on power. *Journal of Experimental Social Psychology*, *49*, 965–971. doi:10.1016/j.jesp.2013.06.001

Peper, E., Booiman, A., Lin, I.-M., & Harvey, R. (2016). Increase strength and mood with posture. *Biofeedback*, *44*, 66–72. doi:10.5298/1081-5937-44.2.04

Ranehill, E., Dreber, A., Johannesson, M., Leiberg, S., Sul, S., & Weber, R. A. (2015). Assessing the robustness of power posing: No effect on hormones and risk tolerance in a large sample of men and women. *Psychological Science*, *26*, 653–656. doi:10.1177/0956797614553946

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172. doi:10.1037/0033-295X.110.1.145

Simmons, J. P., & Simonsohn, U. (2015, May 8). Power posing: Reassessing the evidence behind the most popular TED talk. *Data Colada*. Retrieved from http://web.archive.org/web/20160613100317/http://datacolada.org/37

Simmons, J. P., & Simonsohn, U. (2017). Power posing: *P*-curving the evidence. *Psychological Science*, *28*, 687–693. doi:10.1177/0956797616658563

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). *P*-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547. doi:10.1037/a0033242

Simonsohn, U., Nelson, L., & Simmons, J. (2015). Official user-guide to the *P*-curve. Retrieved from http://www.p-curve.com/guide.pdf

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better *P*-curves: Making *P*-curve analysis more robust to errors, fraud, and ambitious *P*-hacking, a reply to Ulrich and Miller (2015). *Journal of Experimental Psychology: General*, *144*, 1146–1152. doi:10.1037/xge0000104

Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *20*, 57–59.

Wegner, D. M., & Gilbert, D. T. (2000). Social psychology—the science of human experience. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 1–9). Philadelphia, PA: Psychology Press.

Zajonc, R. B. (1998). Emotions. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 591–634). Boston, MA: McGraw-Hill.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*, 75–98. doi:10.1007/s10683-009-9230-z