



# [37] Power Posing: Reassessing The Evidence Behind The Most Popular TED Talk

JOE & URI

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:  
[us19066@gmail.com](mailto:us19066@gmail.com)

DATE RECEIVED:  
July 02, 2015

DOI:  
[10.15200/winn.143585.57764](https://doi.org/10.15200/winn.143585.57764)

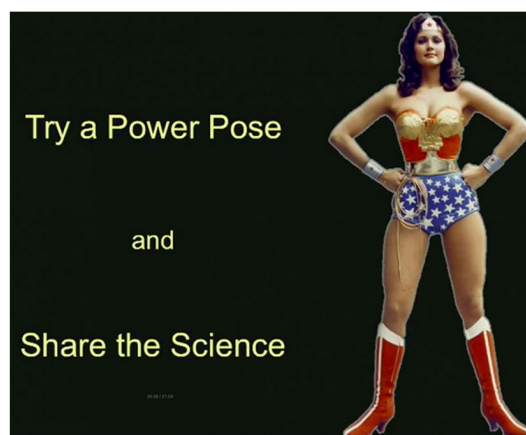
ARCHIVED:  
July 02, 2015

CITATION:  
Joe & Uri, [37] Power Posing: Reassessing The Evidence Behind The Most Popular TED Talk, *The Winnower* 2:e143585.57764, 2015, DOI: [10.15200/winn.143585.57764](https://doi.org/10.15200/winn.143585.57764)

© Uri This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



A recent paper in *Psych Science* ([.pdf](#)) reports a failure to replicate the study that inspired a TED Talk that has been seen 25 million times.<sup>1</sup> The talk invited viewers to do better in life by assuming high-power poses, just like Wonder Woman's below, but the replication found that power-posing was inconsequential.



**TED** Ideas worth spreading

If an original finding is a false positive then its replication is likely to fail, but a failed replication need not imply that the original was a false positive. In this post we try to figure out why the replication failed.

## Original study

Participants in the original study took an expansive “high-power” pose, or a contractive “low-power” pose (*Psych Science* 2010, [.pdf](#)).



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.

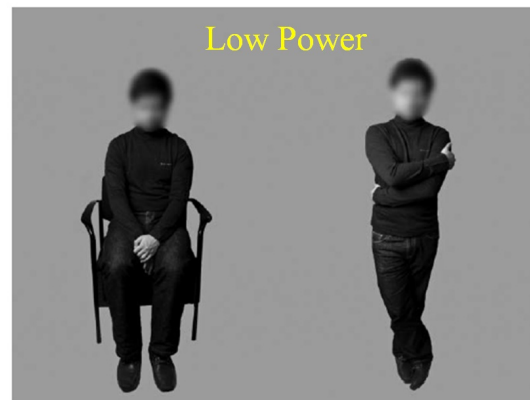


Fig. 2. The two low-power poses used in the study. Participants in the low-power-pose condition were posed in contractive positions with closed limbs.

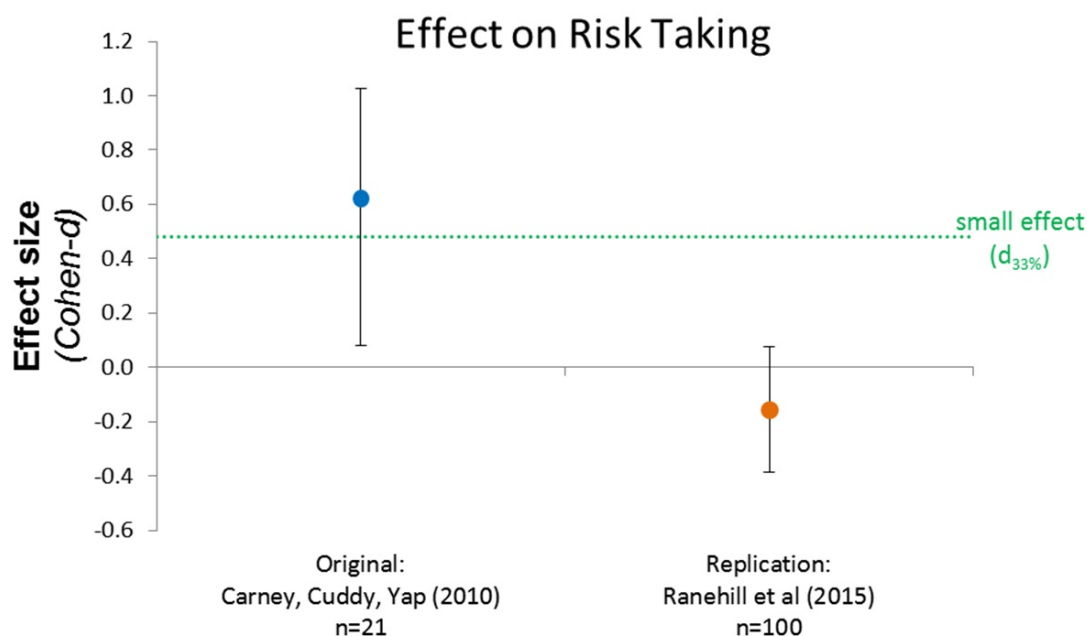
The power-posing participants were reported to have felt more powerful, sought more risk, had higher testosterone levels, and lower cortisol levels. In the replication, power posing affected self-reported power (the manipulation check), but did not impact behavior or hormonal levels.<sup>2</sup> The key point of the TED Talk, that power poses “can significantly change the outcomes of your life” (minute 20:10; [video](#); transcript [.html](#)), was not supported.

### Was The Replication Sufficiently Precise?

Whenever a replication fails, it is important to assess how precise the estimates are. A noisy estimate may be consistent with no effect (i.e., not significant) but also consistent with effects large enough so as to not challenge the original study.

Only precisely estimated non-significant results contradict an original finding (see Colada<sup>[7]</sup> for an example of a strikingly imprecise replication).

Here are the original and replication confidence intervals for the risk-taking measure<sup>3</sup>

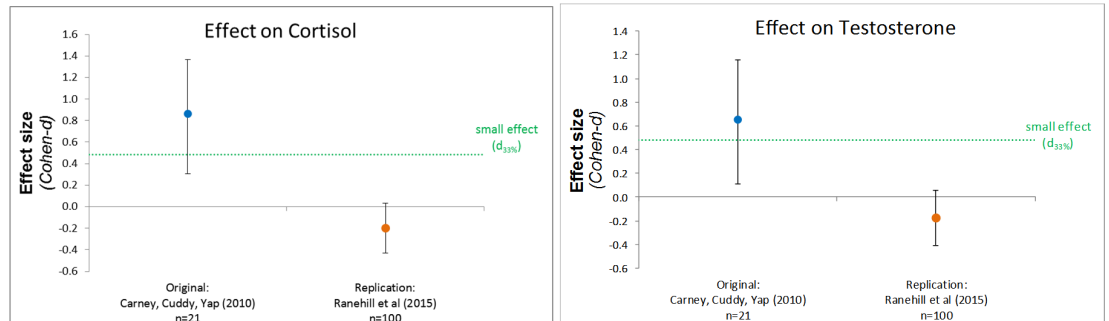


This figure shows that the replication is precise enough to be informative.

The higher end of the confidence interval is  $d = .06$ , meaning that the data are consistent with zero and inconsistent with power-poses affecting risk-taking by more than 6% of a standard deviation. We can put in perspective how small that upper bound is by realizing that, with just  $n = 21$  per cell, the original study would have a meager 5.6% of chance of detecting it (i.e., <6% statistical power).

Thus, even if the effect existed, the replication suggests the original experiment could not have meaningfully studied it. (For more on this approach to thinking about replications check out Uri's Small Telescopes paper ([.pdf](#))).

The same is true for the effects of power poses on hormones:



This implies that there is either a moderator or the original is a false-positive.

### Moderators

In their response ([.pdf](#)), the original authors reviewed every *published* study on power posing they could locate (they found 33), seeking potential moderators, factors that may affect whether power posing works. For example, they speculated that power posing may not have worked in the replication because participants were told that power poses could influence behavior. (One apparent implication of this hypothesis is that watching the power poses TED talk would make power poses ineffective).

The authors also list all differences in design they noted between the original and replication studies (see their nice summary [.pdf](#)).

One approach would be to run a series of studies to systematically manipulate these hypothesized moderators to see whether they matter.

But before spending valuable resources on that, it is necessary to first establish whether there is reason to believe, based on the published literature, that power posing is ever effective. Might it be instead the case that the original findings are false-positive?<sup>4</sup>

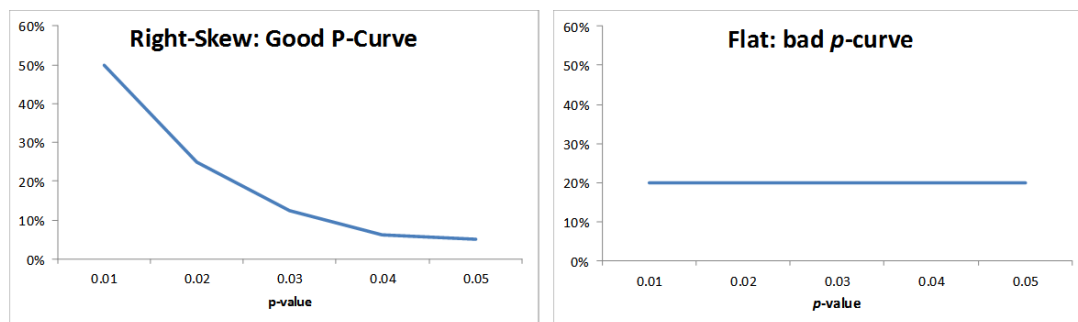
### P-curve

It may seem that 30-some successful studies is enough to conclude the effect is real. However, we need to account for selective reporting. If studies only get published when they show an effect, the fact that all the published evidence shows an effect is not diagnostic.

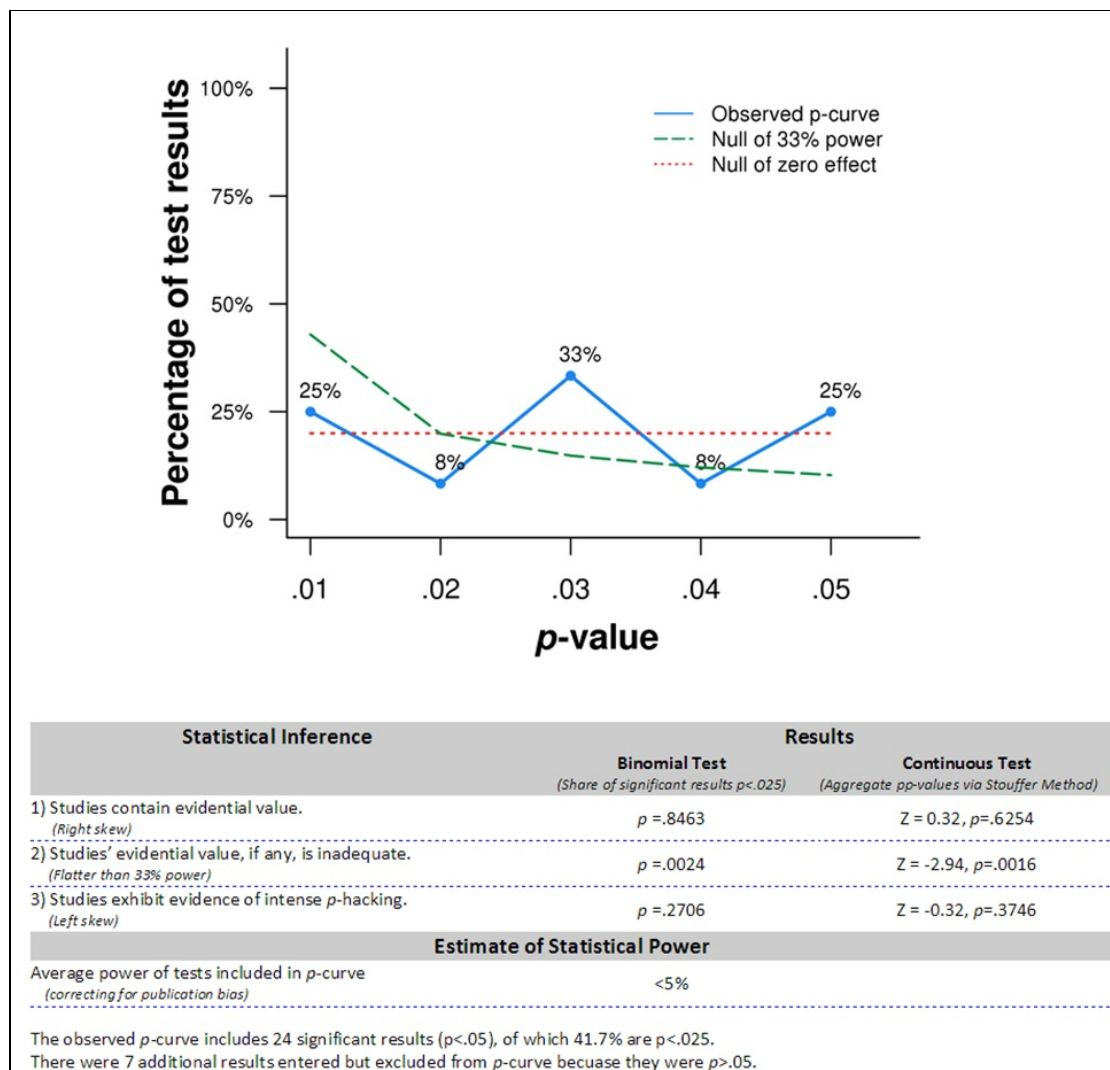
P-curve is just the tool for this. It tells us whether we can rule out selective reporting as the sole explanation for a set of  $p < .05$  findings (see *p-curve* paper [.pdf](#)).

We conducted a p-curve analysis on the 33 studies the original authors cited as evidence that power posing works in their reply. They come from many labs around the world.

If power posing were a true effect, we should see a curve that is right-skewed, tall on the left and low on the right. If power posing had not effect, we expect a flat curve.



The actual  $p$ -curve for power posing, is flattish, definitely not right-skewed.



Note: you should ignore  $p$ -curve results that do not include a disclosure table; here is ours ([xlsx](#)) .

Consistent with the replication motivating this post,  $p$ -curve indicates that either power-posing overall has no effect, or the effect is too small for the existing samples to have meaningfully studied it. Note that there are perfectly benign explanations for this: e.g., labs that run studies that worked wrote them up, labs that run studies that didn't, didn't.<sup>5</sup>

While the simplest explanation is that all studied effects are zero, it may be that one or two of them are real (any more and we would see a right-skewed  $p$ -curve). However, at this point the evidence for the basic effect seems too fragile to search for moderators or to advocate for people to engage in power posing to better their lives.



#### Author feedback.

We shared an early draft of this post with the authors of the original and failed replication studies. We also shared it with Joe Cesario, author of a few power-posing studies with whom we had discussed this literature a few months ago.

Note: It is our policy not to comment, they get the last word

**Amy Cuddy** ([.html](#)), co-author of the original study and TED talk speaker, provided useful suggestions for clarifications and modification that lead to several revisions, including a new title. She also sent this note:

I'm pleased that people are interested in discussing the research on the effects of adopting expansive postures. I hope, as always, that this discussion will help to deepen our understanding of this and related phenomena, and clarify directions for future research. Given that our quite thorough response to the Ranehill et al. study has already been published in *Psychological Science*, I will direct you to that, ([.html](#)). I respectfully disagree with the interpretations and conclusions of Simonsohn et al., but I'm considering these issues very carefully and look forward to further progress on this important topic.

**Roberto Weber** ([.html](#)), co-author of the failed replication study sent Uri a note, co-signed with all co-authors, which he allowed us to post (see it here [.pdf](#)). They make four points, we quote attempting to summarize:

- (1) none of the 33 published studies, other than the original and our replication, study the effect [...] on hormones
- (2) even within [the original] the hormone result seems to be presented inconsistently
- (3) regarding differences in designs [from Table 2 in their response], [...] the evidence does not support many of these specific examples as probable moderators in this instance
- (4) we employed an experimenter-blind design [...] this might be the factor that most plausibly serves as a moderator of the effect"

Again, full response: [.pdf](#).

**Joe Cesario** ([html](#))

Excellent post, very clear and concise. Our lab has also been concerned with some of the power pose claims—less about direct replicability and more about the effectiveness in more realistic situations (motivating our Cesario & McDonald paper and another paper in prep with David Johnson). These concerns also led us to contact a number of the most cited power pose researchers to invite them for a multi-site, collaborative replication and extension project. Several said they could not commit the resources to such an endeavor, but several did agree in principle. Although progress has stalled, we are hopeful about such a project moving forward. This would be a useful and efficient way of providing information about replicability as well as potential moderation of the effects.

#### SUBSCRIBE TO BLOG VIA EMAIL

Enter your email address to subscribe to this blog and receive notifications of new posts by email.

Email Address

### Footnotes.

1. This TED talk is ranked 2nd in overall downloads and 1st in per-year downloads. The talk also shared a very touching personal story of how the speaker overcame an immense personal challenge; this post is concerned exclusively with the science part of the talk [↩]
2. The original authors consider “feelings of power” to be a manipulation check rather than an outcome measure. In their most recent paper they write, “as a manipulation check, participants reported how dominant, in control, in charge, powerful, and like a leader they felt on a 5-point scale” (.pdf) [↩]
3. The definition of Small Effect in the figure corresponds to an effect that would give the original sample size 33% power, see Uri’s Small Telescopes paper (.pdf) [↩]
4. \*\*\*See Simine Vazire’s excellent post on why it’s wrong to always ask authors to explain failures to replicate by proposing moderators rather than concluding that the original is a false-positive .html (the \*\*\* is a gift for Simine) [↩]
5. Because the replication obtained a significant effect of power posing on the manipulation check, self-reported power, we constructed a separate  $p$ -curve including only the 7 manipulation check results. The resulting  $p$ -curve was directionally right-skewed ( $p=.075$ ). We interpret this as further consistency between  $p$ -curve and the replication results. If from studies reporting both manipulation checks and the effect of interest we select only the manipulation check, and from other studies we select the effect of interest (something we believe is not reasonable but report anyway in case a reader disagrees and for the sake of robustness), the overall conclusions do not change. The overall  $p$ -curve still looks flat, does not conclude there is evidential value ( $Z=.84, p=.199$ ), and does conclude the curve is flatter than 33% ( $Z=2.05, p=.0203$ ) [↩]