Lecture #11

Current Issues: The power pose debate

Questionable Research Practices (QRPs)

Research Report

ASSOCIATION FOR
PSYCHOLOGICAL SCIENCE

Psychological Science
21(10) 1363–1368
© The Author(s) 2010
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797610383437
http://pss.sagepub.com
SAGE

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.

Questions

- Can enacting a powerful posture cause a person to become more powerful?

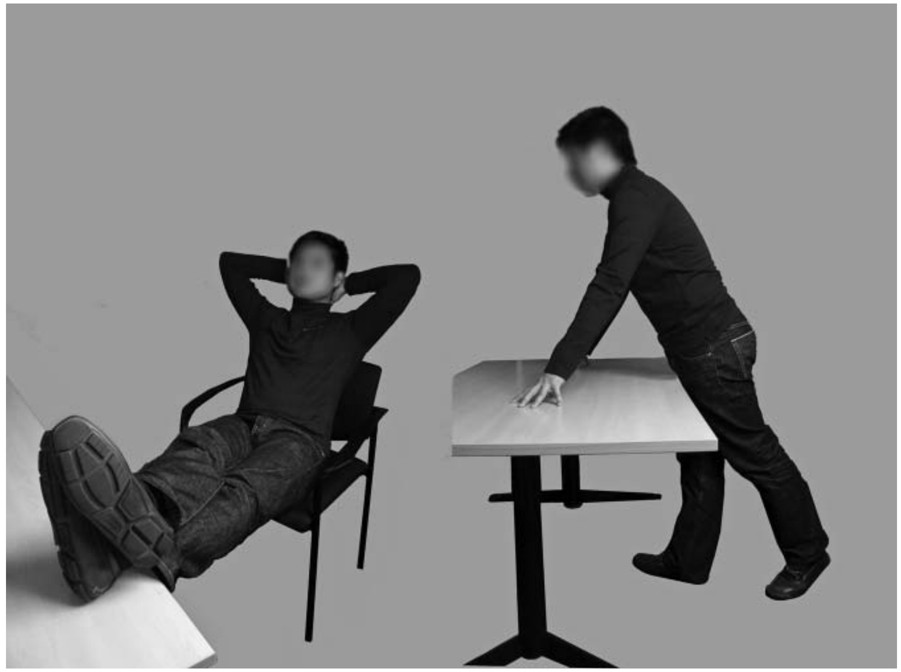


Fig. 1. The two high-power poses used in the study. Participants in the high-power-pose condition were posed in expansive positions with open limbs.

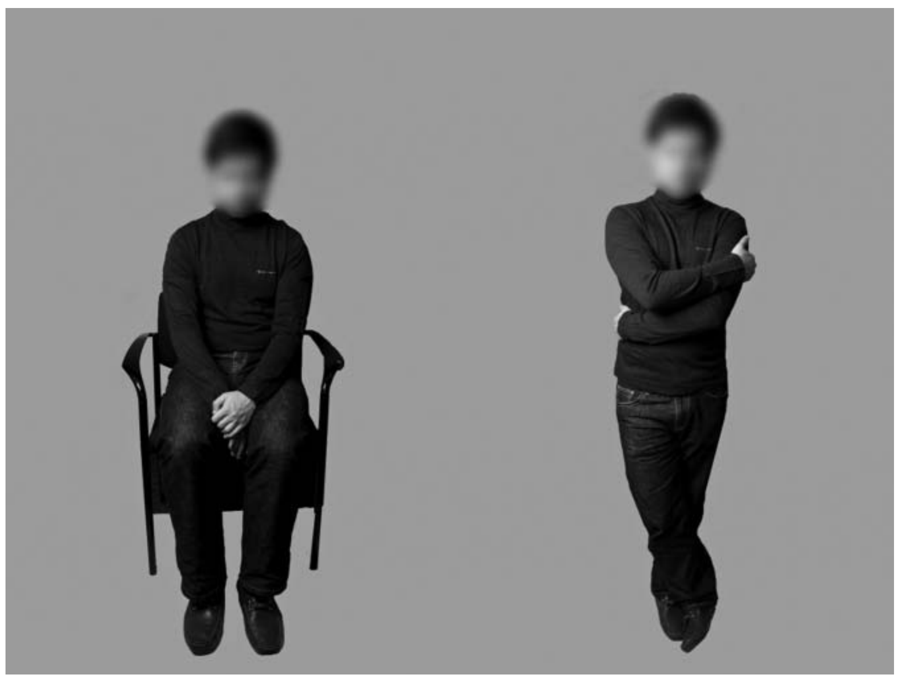


Fig. 2. The two low-power poses used in the study. Participants in the low-power-pose condition were posed in contractive positions with closed limbs.

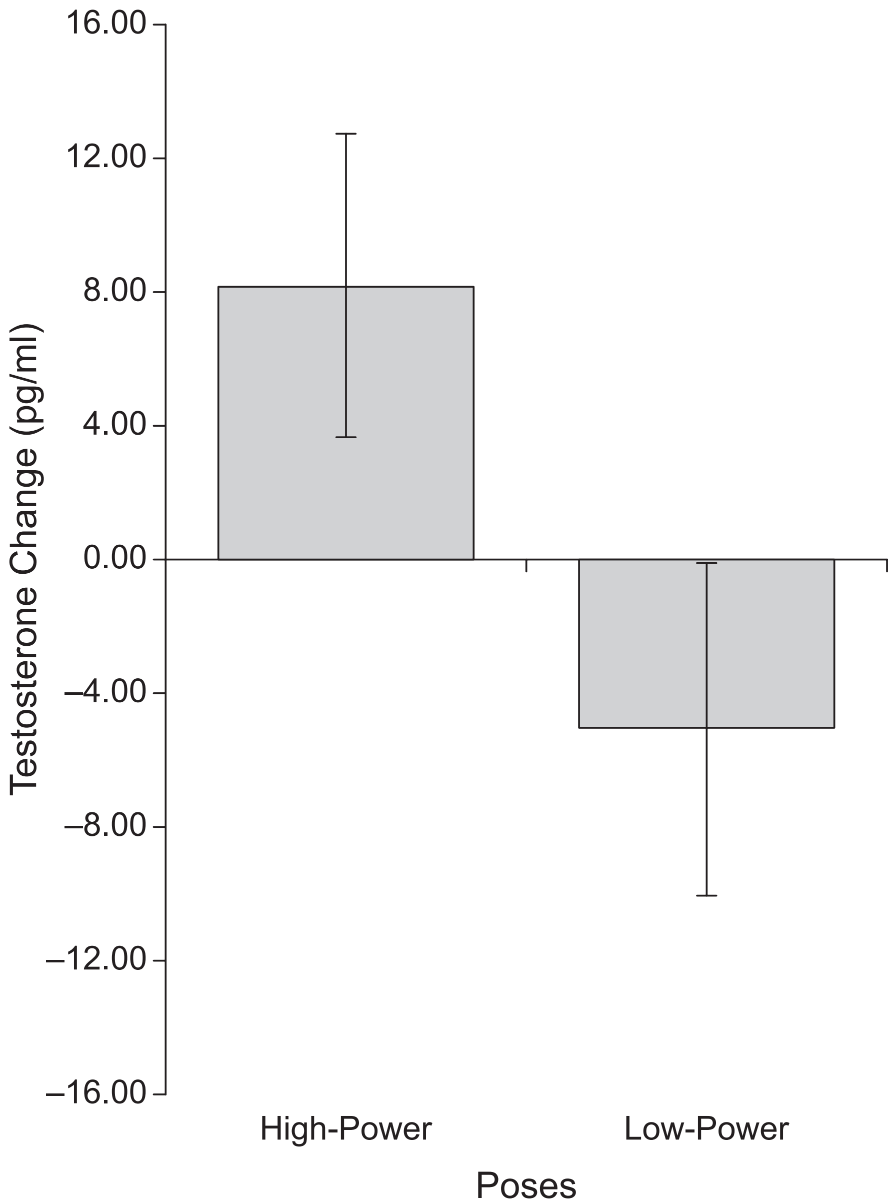


Fig. 3. Mean changes in the dominance hormone testosterone following high-power and low-power poses. Changes are depicted as difference scores (Time 2 – Time 1). Error bars represent standard errors of the mean.

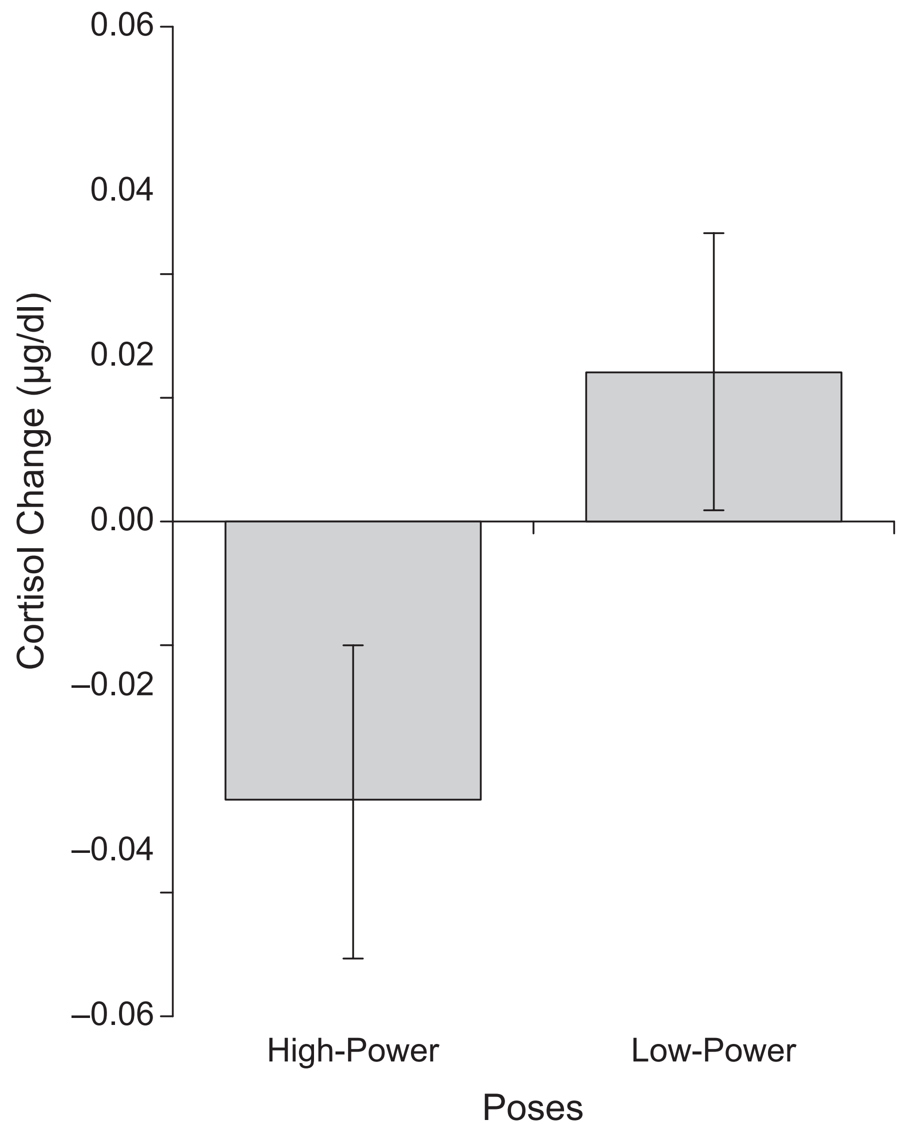


Fig. 4. Mean changes in the stress hormone cortisol following high-power and low-power poses. Changes are depicted as difference scores (Time 2 – Time 1). Error bars represent standard errors of the mean.

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.

Conclusions

- “By simply changing physical posture, an individual pre- pares 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.”

2012 Ted Talk

- https://www.ted.com/talks/amy_cuddy_your_body_language_shapes_who_you_are

Around the same time

- Many notable failures to replicate priming type effects
- High-profile cases of data-fraud
- A growing sense that the field needs to address these problems
- http://www.nature.com/news/replication-studies-bad-copy-1.10634

Replication Failures

- Doyen et al failed to replicate Bargh’s classic aging priming study (discussed several lectures ago).

Silly Science

- Bem reported a series of experiments purportedly showing evidence of “psychic” or “psi” phenomena.
- 3 labs could not replicate the results
- Led some to wonder how many other “results” would fail to replicate

Data Fraud

- 2011 - Diederik Stapel, high-profile social psychologist, published in many top journals including Science
  - Fabricated data for over 30+ publications
  - http://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?pagewanted=all

Kahneman’s open letter

- https://www.nature.com/news/nobel-laureate-challenges-psychologists-to-clean-up-their-act-1.11535

The Open Science Replication project

- Discussed in First class
- Many labs attempted to replicate many existing findings across many journals
- Many findings did not replicate

Back to Power posing

Commentary

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
1–4
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
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

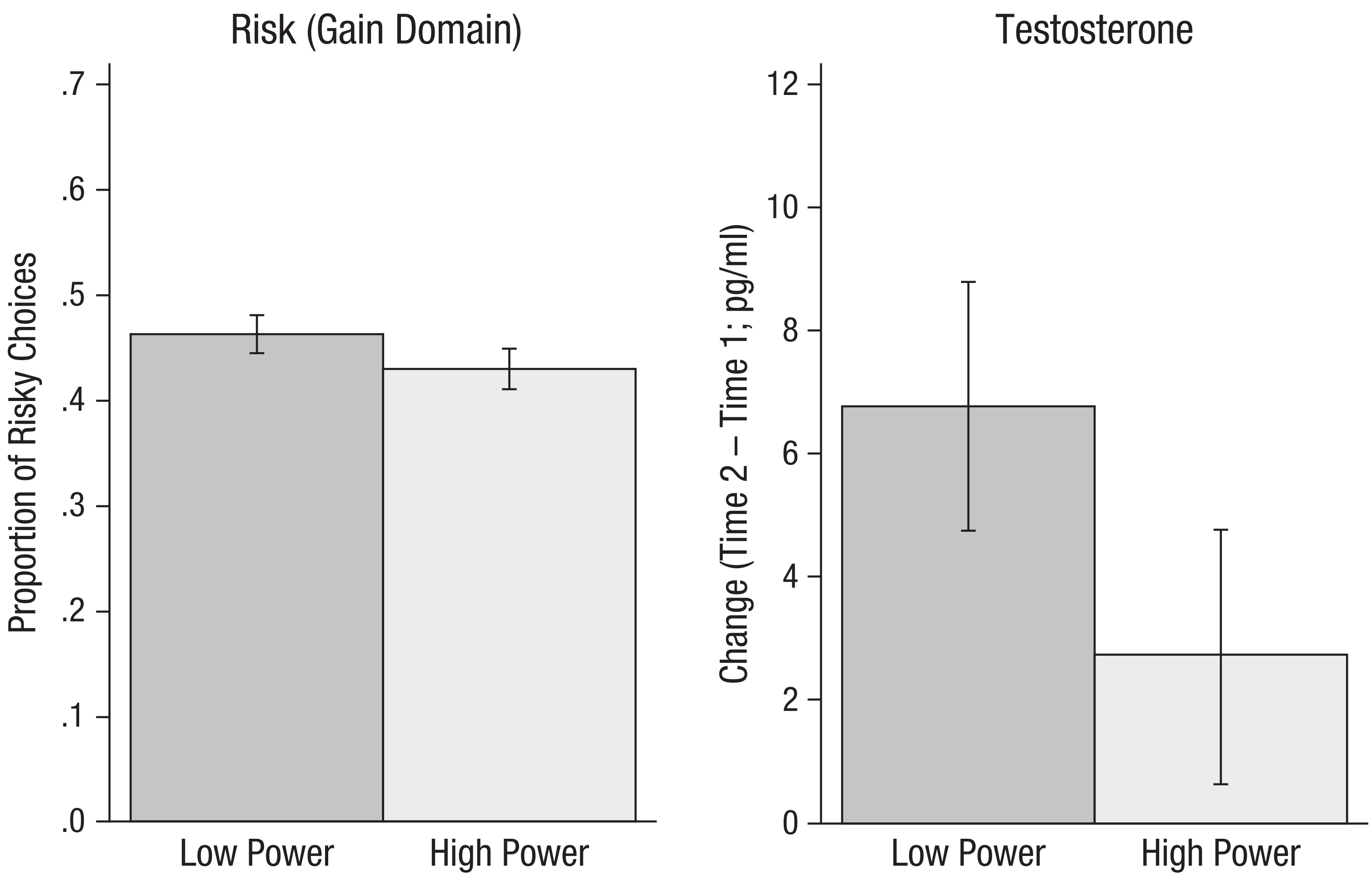


Fig. 1. Mean proportion of risky choices in the gain domain (left) and mean change in testosterone from before the power-pose manipulation (Time 1) to 17 min after the power-pose manipulation (Time 2; right). For each graph, results are shown separately for the high- and low-power-pose conditions. Error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

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.

Commentary

Psychological Science
1–7
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797614566855
pss.sagepub.com

Review and Summary of Research on the Embodied Effects of Expansive (vs. Contractive) Nonverbal Displays

Dana R. Carney[1], Amy J. C. Cuddy[2], and Andy J. Yap[3]
[1]Department of Psychology, University of California, Berkeley; [2]Harvard Business School, Harvard University; and [3]Department of Organisational Behaviour, INSEAD

Received 9/10/14; Revision accepted 12/12/14

Carney's (first author) new position on power pose effects

- http://faculty.haas.berkeley.edu/dana_carney/pdf_My%20position%20on%20power%20poses.pdf

Cuddy article 1

- https://www.thecut.com/2016/09/read-amy-cuddys-response-to-power-posing-critiques.html

Recent background article

- https://www.nytimes.com/2017/10/18/magazine/when-the-revolution-came-for-amy-cuddy.html

QRPs (questionable research practices)

Research Article

ASSOCIATION FOR
PSYCHOLOGICAL SCIENCE

Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling

Psychological Science
23(5) 524–532
© The Author(s) 2012
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797611430953
http://pss.sagepub.com
SAGE

Leslie K. John[1], George Loewenstein[2], and Drazen Prelec[3]
[1]Marketing Unit, Harvard Business School; [2]Department of Social & Decision Sciences, Carnegie Mellon University; and [3]Sloan School of Management and Departments of Economics and Brain & Cognitive Sciences, Massachusetts Institute of Technology

Abstract

Cases of clear scientific misconduct have received significant media attention recently, but less flagrantly questionable research practices may be more prevalent and, ultimately, more damaging to the academic enterprise. Using an anonymous elicitation format supplemented by incentives for honest reporting, we surveyed over 2,000 psychologists about their involvement in questionable research practices. The impact of truth-telling incentives on self-admissions of questionable research practices was positive, and this impact was greater for practices that respondents judged to be less defensible. Combining three different estimation methods, we found that the percentage of respondents who have engaged in questionable practices was surprisingly high. This finding suggests that some questionable practices may constitute the prevailing research norm.

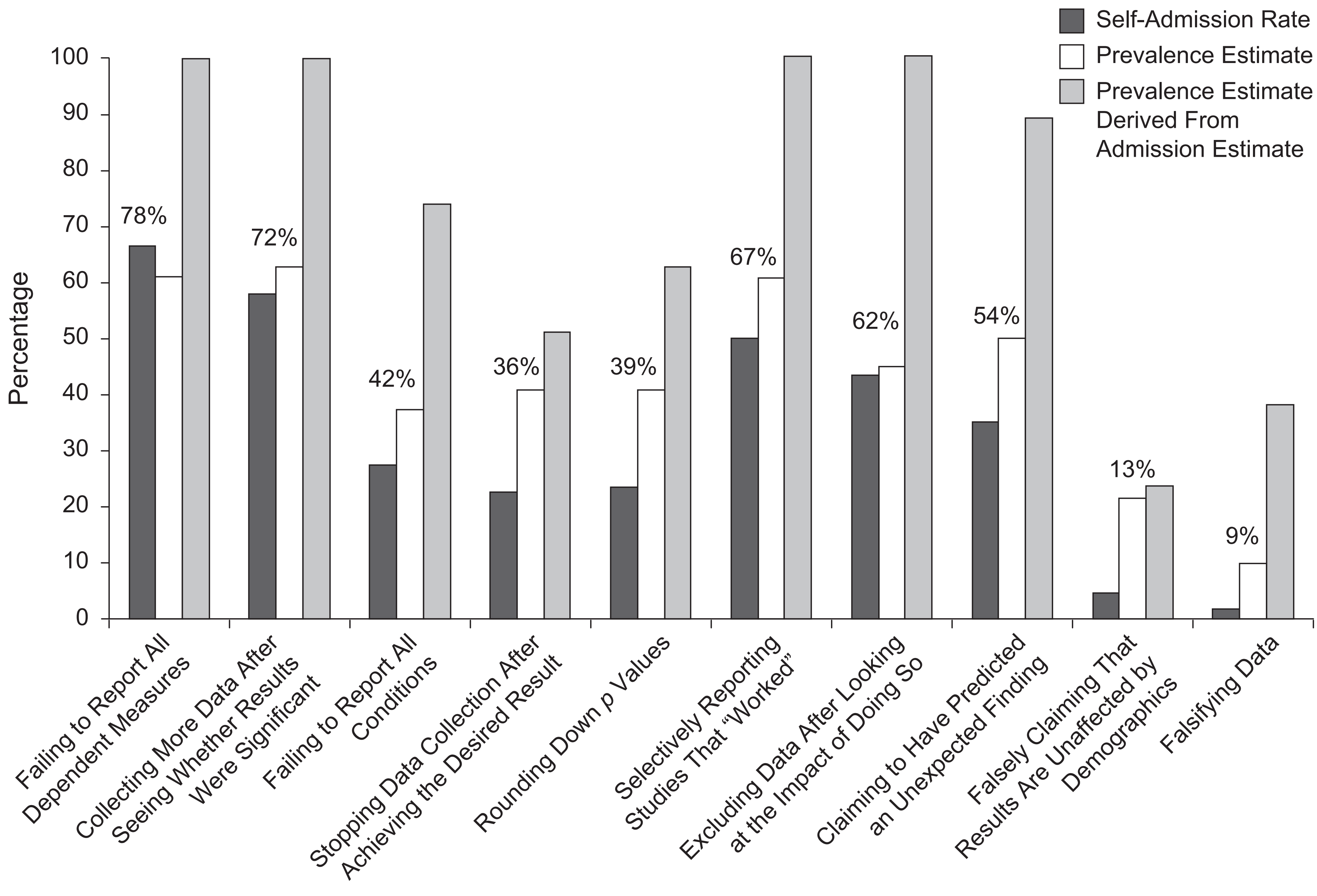


Fig. 1. Results of the Bayesian-truth-serum condition in the main study. For each of the 10 items, the graph shows the self-admission rate, prevalence estimate, prevalence estimate derived from the admission estimate (i.e., self-admission rate/admission estimate), and geometric mean of these three percentages (numbers above the bars). See Table 1 for the complete text of the items.

1. Failing to report all DVs in a study

1. Failing to report all DVs in a study

ReportedDVs	Hypothetical Unreported DVs	working memory
Feelings of power	Heart rate	IQ
Risky choice	pupil dilation	other questionnaires about power
Testosterone	other biomarkers	other behavioral measures…

1. Failing to report all DVs in a study

- Can be questionable when:
  - measures of the same construct show different patterns of significance, and non-significant results are omitted
  - most the DVs are exploratory and the non-significant ones are omitted

2. Collecting more data after seeing whether the results were significant

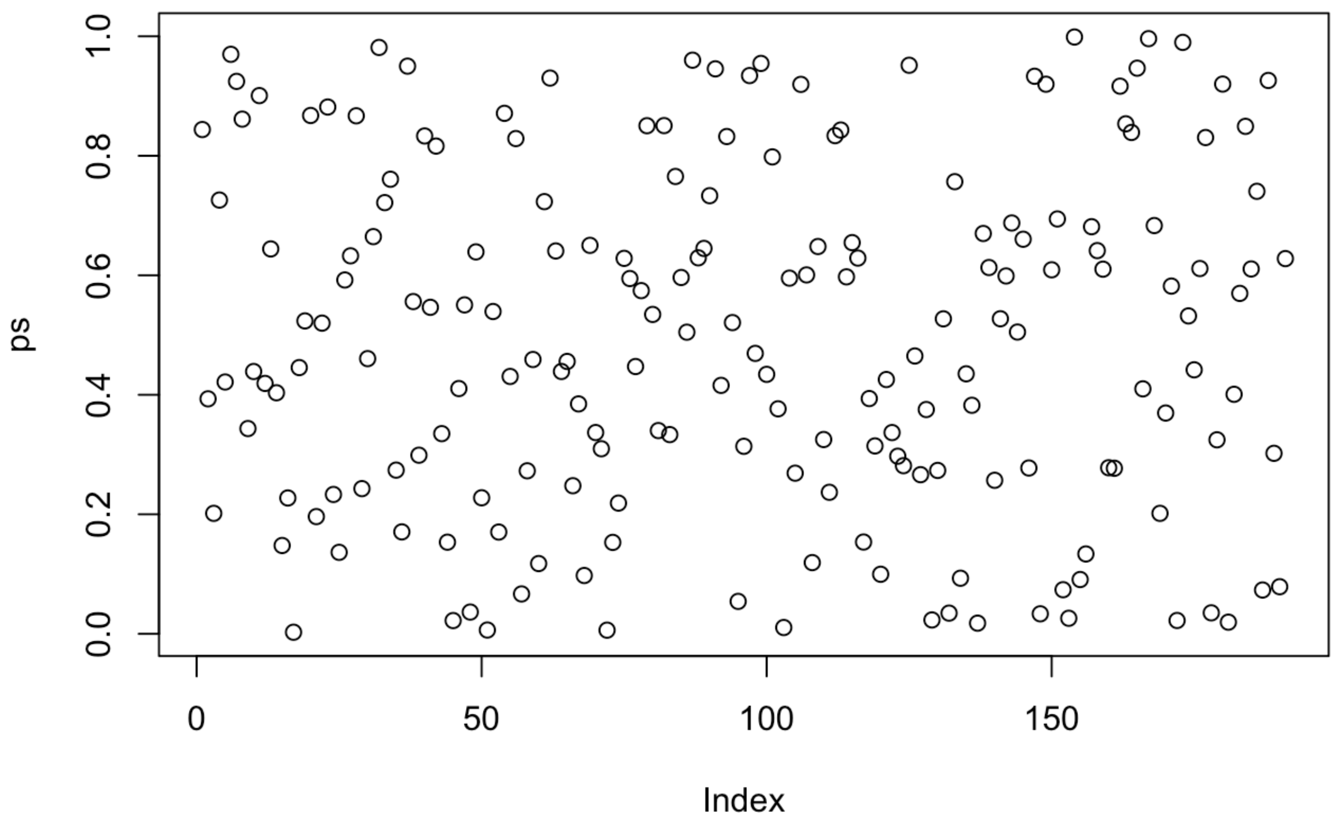
P-Hacking

- Researchers have many options to “massage” their data into reaching significance
- P-hacking occurs when researchers engage in practices (knowingly or unknowingly) that increase the chances of finding a significant result

P-Hacking 1: Increasing N

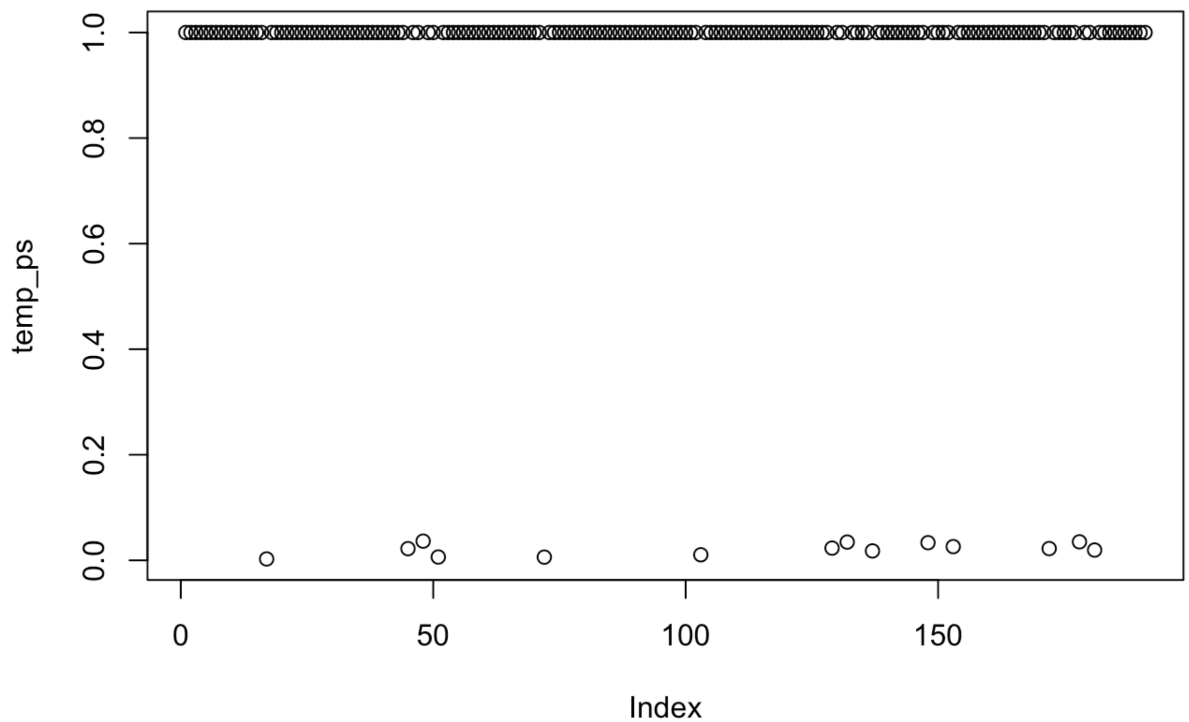
- Increasing sample-size increases power to detect an effect when it is there
- But, conducting new statistical tests every time new subjects are added increases Type I error (False-positive, incorrectly rejecting the null, or finding a significant result due to chance)

Simulation of P-values from a one-sample test computed every time a new subject is added (null is assumed)



Index is number of subjects

Simulation of P-values from a one-sample test computed every time a new subject is added (null is assumed)



Same as before, but all p > .05 are set to 1
We see the null is rejected 14 times over the course of adding additional subjects

Best Practice

- Set N in advance of the study based on a plausible power analysis

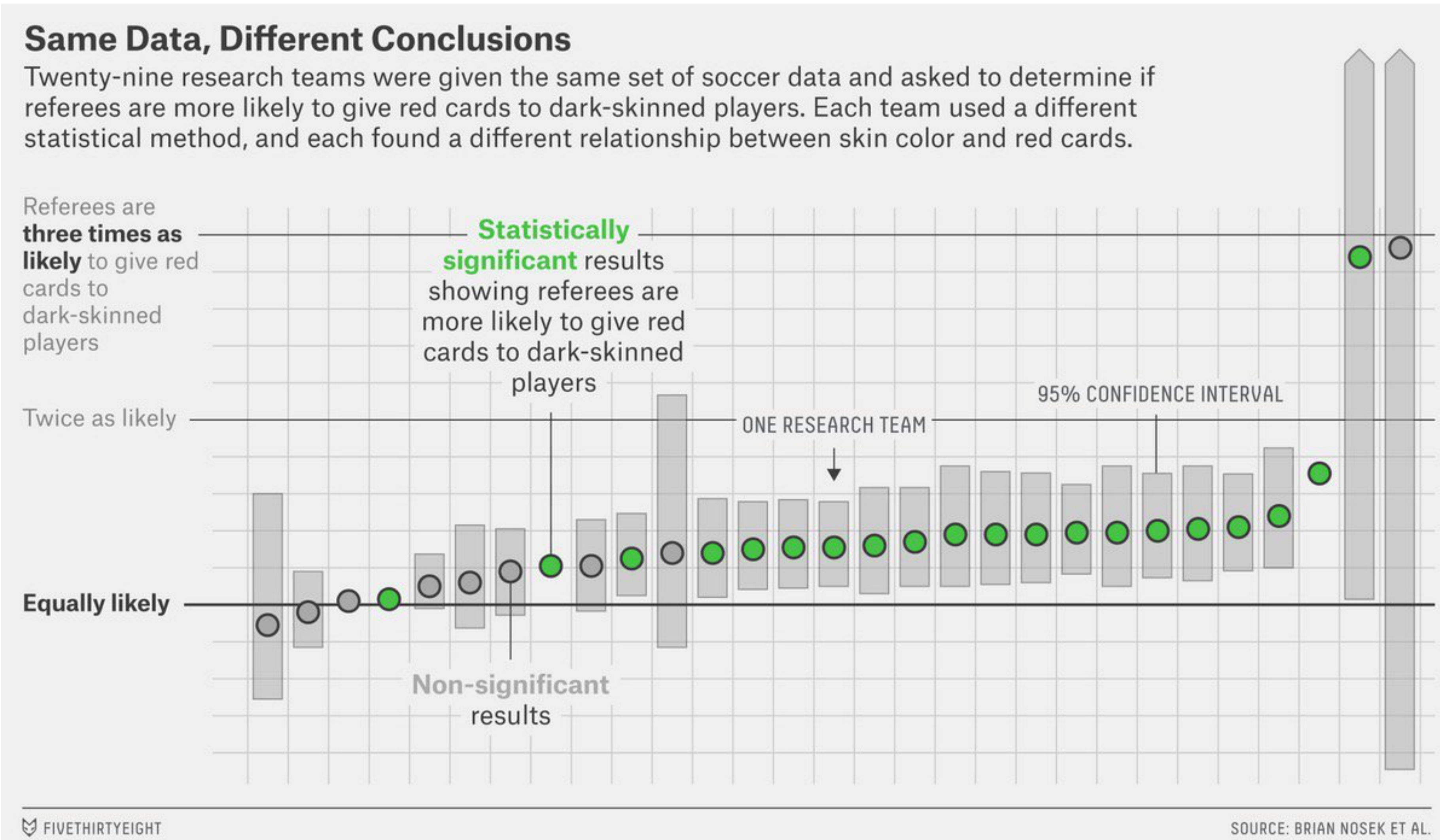
- Do not incrementally add subjects and re-test for significance

- If adding more N, then conduct a complete replication

P-Hacking 2: Trying different tests

- A data-set can usually be analyzed in many different ways, with different tests that produce slightly different results
- A researcher could inflate their type I error rate by conducting many different kinds of tests on the same data, and then reporting only the ones that are significant

Same data, different researchers produce different results



Best Practice

- Plan and justify your statistical tests prior to data collection
  - Create simulated data and analyze with your planned tests to demonstrate that they are appropriate to answer to your question

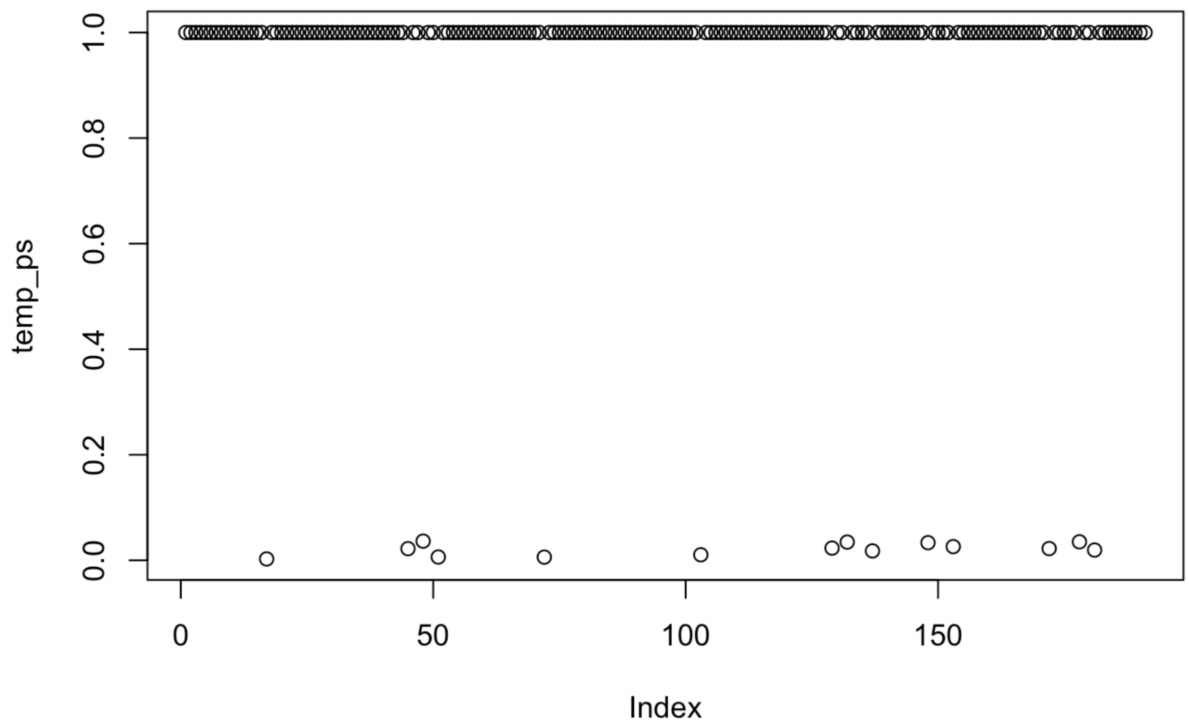
3. Failing to report all conditions

3. Failing to report all conditions

- Questionable when:
  - Data in other conditions are omitted because the make the story less clear

4. Stopping data collection after reaching significance

Simulation of P-values from a one-sample test computed every time a new subject is added (null is assumed)



Same as before, but all p > .05 are set to 1
We see the null is rejected 14 times over the course of adding additional subjects

5. Rounding down p-values

- e.g., you find $p=.054$, your report, $p<.05$

6. Selectively reporting studies that worked

File drawer problem

- Studies with non-significant results have lower publication rates (left in the file-drawer)
- Problem: the sizes of published effects could be inflated because null-results are not included in the estimate

7. Excluding data after looking at the impact of doing so

7. Excluding data after looking at the impact of doing so

- e.g., my effect is $p=.06$
    - but, I have one “weird” subject, and if I remove them then the effect is $p<.05$

7. Excluding data after looking at the impact of doing so

- Best Practice
  - 1) determine subject inclusion and exclusion parameters before study
  - 2) Justify exclusion and inclusion parameters
  - 3) apply these parameters uniformly across studies

8. Claiming to have predicted an unexpected finding

HARKING

- Hypothesizing after the results are known

HARKING

Researcher's theory predicts	Exploratory IVs included in study		
Effect of interest	IV2	IV3	IV4
But, Effect is not observed	non-significant	non-significant	Interacts with effect of interest
			HARKING (Researcher Invents new theory after the fact, and claims to have predicted unexpected effect)

9. Falsely claiming that results are unaffected by demographics

10. Falsifying data

Carney's (first author) new position on power pose effects

- http://faculty.haas.berkeley.edu/dana_carney/pdf_My%20position%20on%20power%20poses.pdf