Psyc 60: Intro to Statistics

Prof. Judith Fan

Spring 2021

All remaining due dates

			What you have learned *Before:* Chapter 12		
10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5	Project Milestone 5 Du (Project Report
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster)	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due

All remaining due dates

			What you have learned *Before:* Chapter 12		
10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5	Project Milestone 5 Du (Project Report
					Final

Note: We will be instituting class-wide no-penalty extensions for ALL CourseKata modules. That means that so long as you complete any missing modules by **Friday 11:59PM on June 4**, you will receive full credit. (However, no extensions granted beyond that!)

All remaining due dates

10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	**What you have learned** *Before:* Chapter 12 SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5	Project Milestone 5 Du (Project Report
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster)	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due

All remaining due dates

10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	**What you have learned** *Before:* Chapter 12 SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5		Project Milestone 5 Du (Project Report
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster)	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due	

Released Thursday at 5PM & due by 4:59PM on Friday

All remaining due dates

10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	**What you have learned** *Before:* Chapter 12 SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5	Project Milestone 5 Du (Project Report
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster)	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due

Due by 11:59PM PT on DataHub.

All remaining due dates

			What you have learned *Before:*		
	Memorial		Chapter 12		
			methodological trends in psychology		
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster)	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due

Thank you to everyone who has already submitted a CAPE! :)
We're at 18/85 (~22% of the class) now. If we reach > 80% of the class, everyone's lowest quiz score will be automatically dropped.

All remaining due dates

10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	**What you have learned** *Before:* Chapter 12 SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5	Project Milestone 5 Du (Project Report
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due

Due by 11:59PM PT on DataHub.

All remaining due dates

10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	**What you have learned** *Before:* Chapter 12 SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5	Project Milestone 5 Du (Project Report
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster)	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due

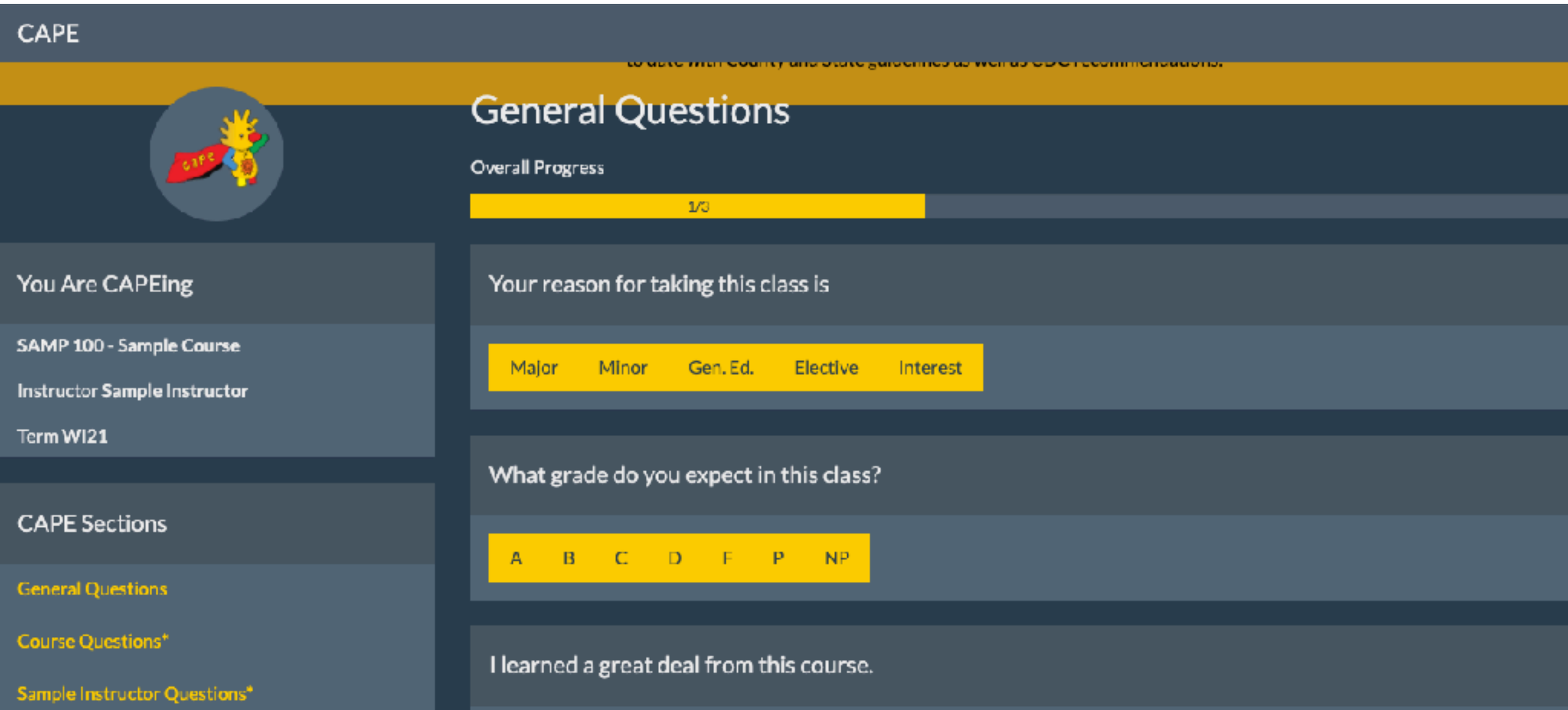
In the same Zoom room we are in right now!

All remaining due dates

10	May 31	**Memorial Day (No class)** *Before:* N/A *During:* N/A	**What you have learned** *Before:* Chapter 12 SONA Credits Due *During:* Broader methodological trends in psychology	Quiz 5	Project Milestone 5 Du (Project Report
11	Jun 7	CAPES Due (before 8am)	Project Milestone 6 Due (Poster)	Final Project Showcase (11:30am-2:30pm)	Lab 5 Due

Due by 11:59PM PT on DataHub.

Every time someone fills out a CAPE, an actual unicorn does a little dance.



Last Time

Lab 5b: Hypothesis Testing

General announcements

Break out into lab groups

Want real-time help?
(1) Post to #lab-assignments, mention both your TA & Room
(2) "Asking for help" in Zoom
(#2 is a bit less reliable!)

Return to main room and debrief

Everyone come back at 2:10pm PT

Today

Lecture 20: broader methodological trends in psychology

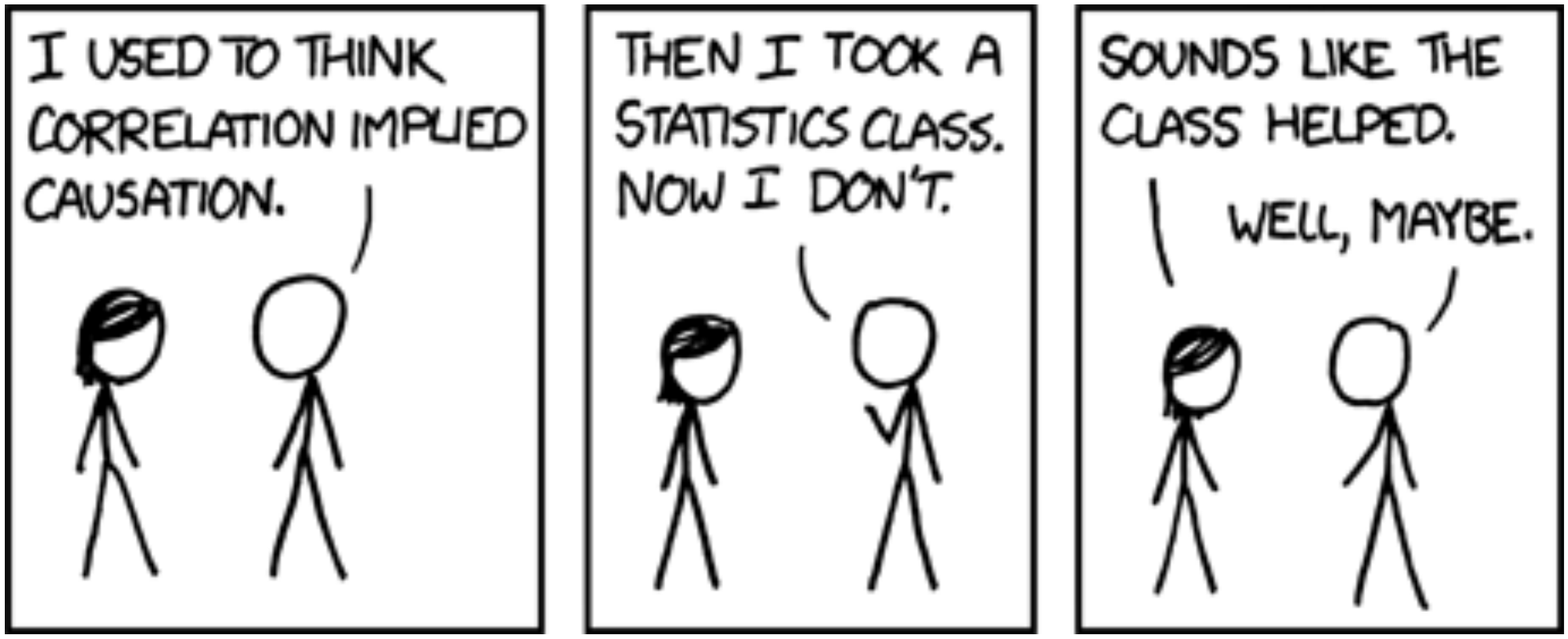
How does correlation relate to causation?

Problems with status quo: p-hacking, HARK-ing

Open science and computational reproducibility

1 *How does correlation relate to causation?*

Correlation and causation



https://xkcd.com/552/

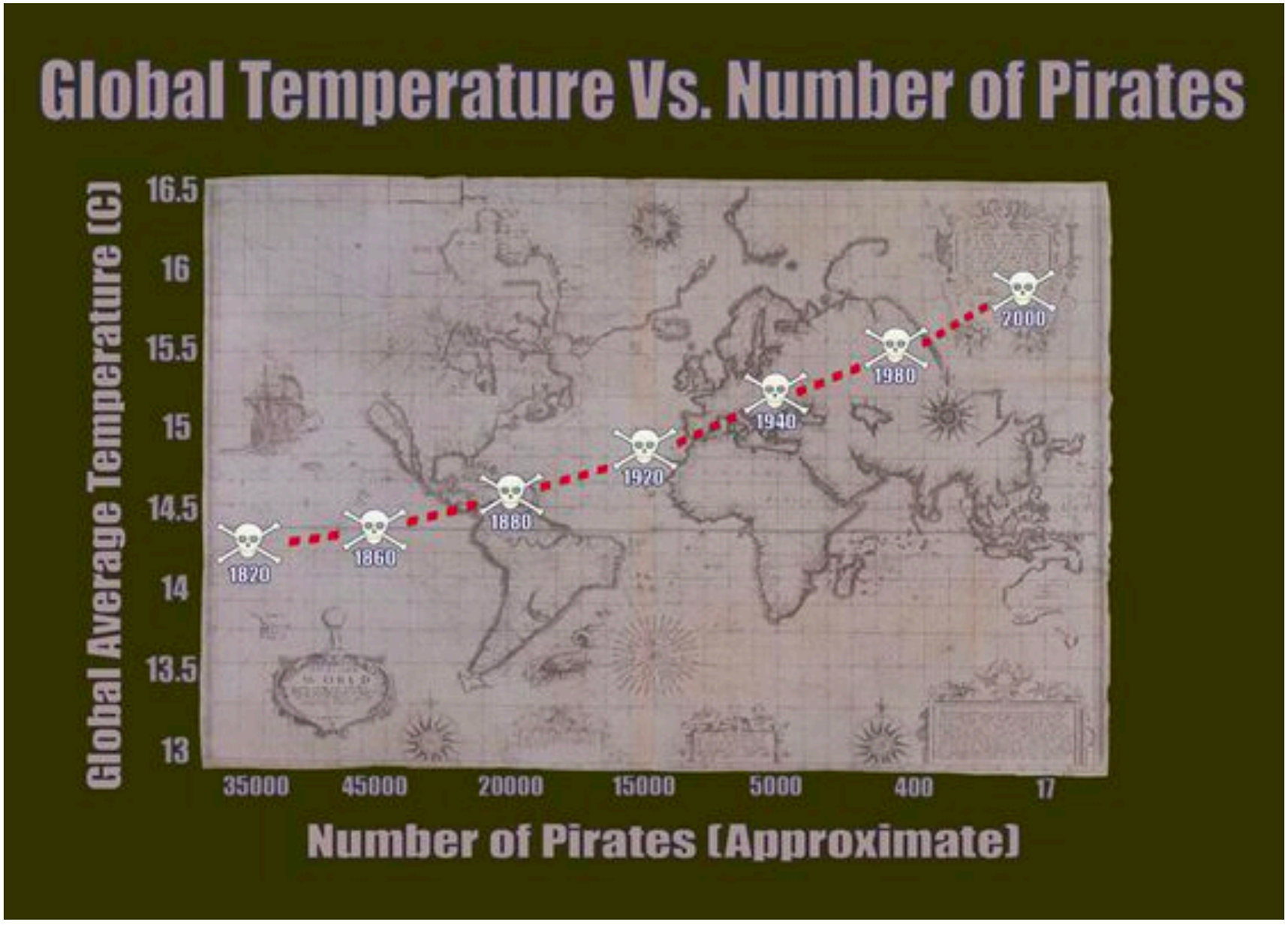
1 *How does correlation relate to causation?*

Correlation and causation

- So many of you might be familiar with the mantra: "correlation doesn't imply causation."
- It's a useful reminder because as human beings, we are extremely good at detecting patterns and finding meaning in the patterns we see.

1 *How does correlation relate to causation?*

Correlation and causation



https://www.forbes.com/sites/erikaandersen/2012/03/23/true-fact-the-lack-of-pirates-is-causing-global-warming/

http://www.tylervigen.com/spurious-correlations

1 *How does correlation relate to causation?*

Correlation and causation

- So many of you might be familiar with the mantra: "correlation doesn't imply causation."
- It's a useful reminder because as human beings, we are extremely good at detecting patterns and finding meaning in the patterns we see.
- How should we think about the relationship between correlation and causation? Oftentimes observational data is the only kind available.

“

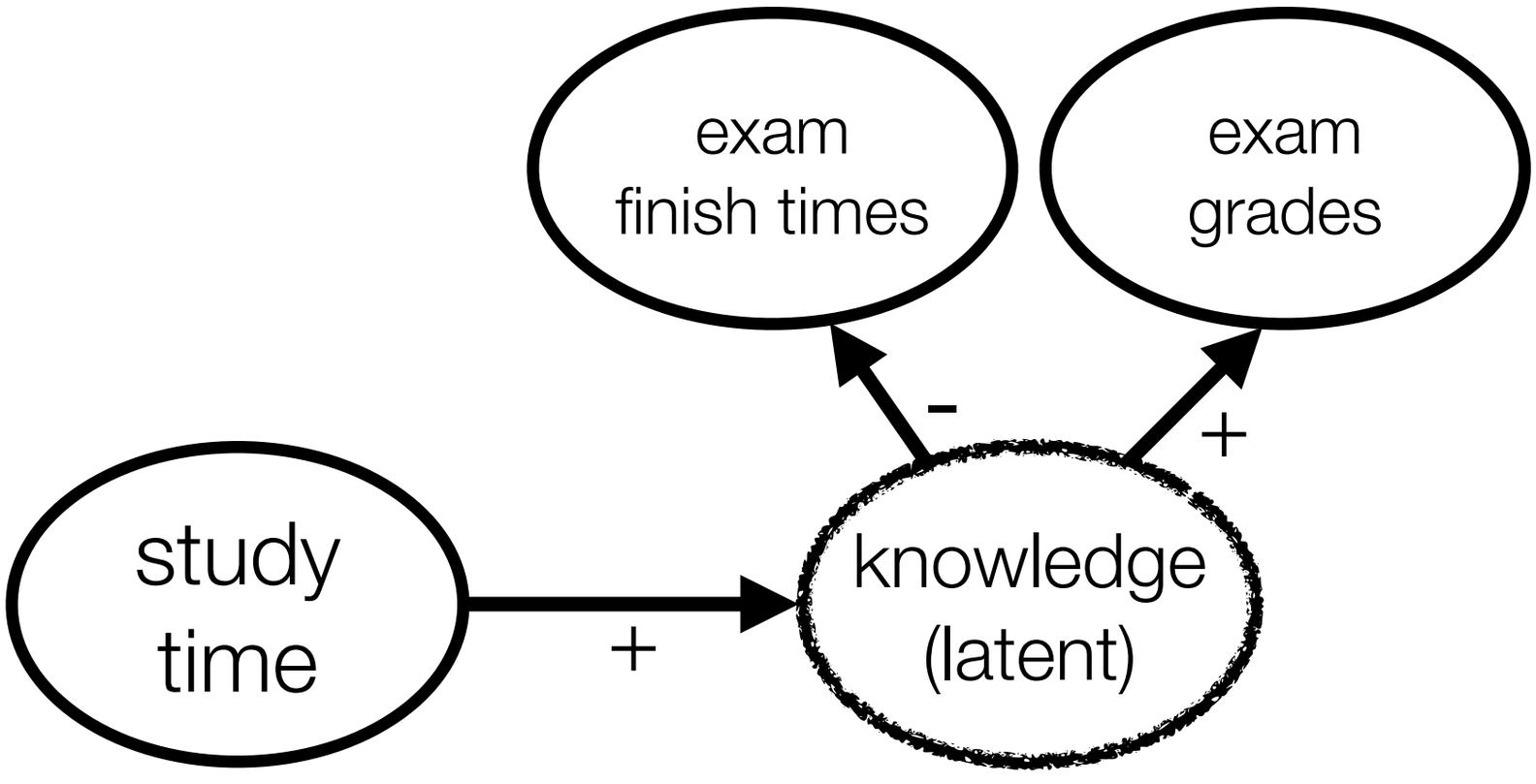
“Correlation does not imply causation, but it’s a pretty good hint.”

-Edward Tufte

1 *How does correlation relate to causation?*

Understanding causation using causal graphs

A causal graph describes the latent causal relations that give rise to the variables that we measure



arrows reflect causal relations

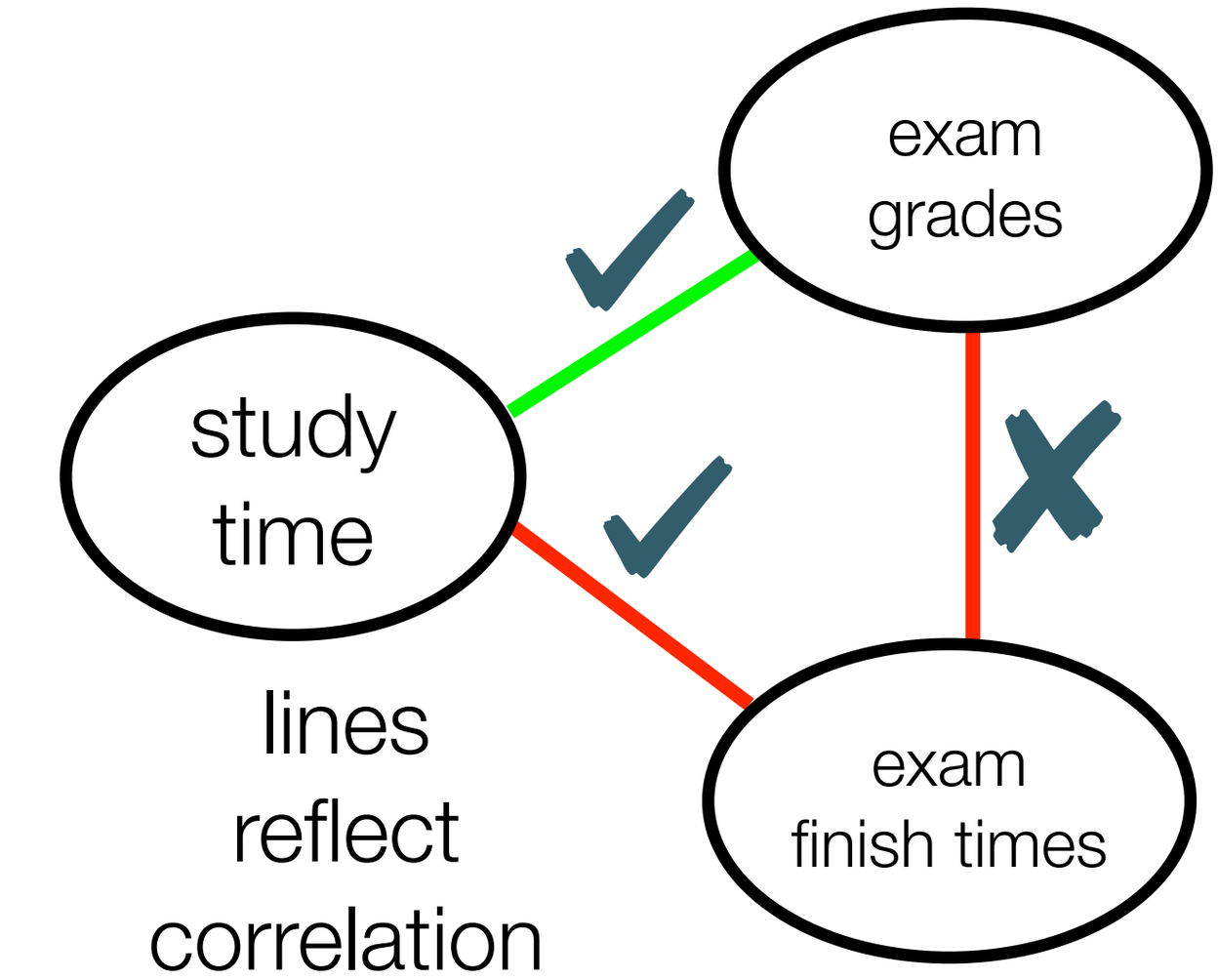
Causal relations mean that manipulating one variable will change another

Increasing study time will increase knowledge, which increases grades and reduces exam finishing time

1 *How does correlation relate to causation?*

Correlation and causation

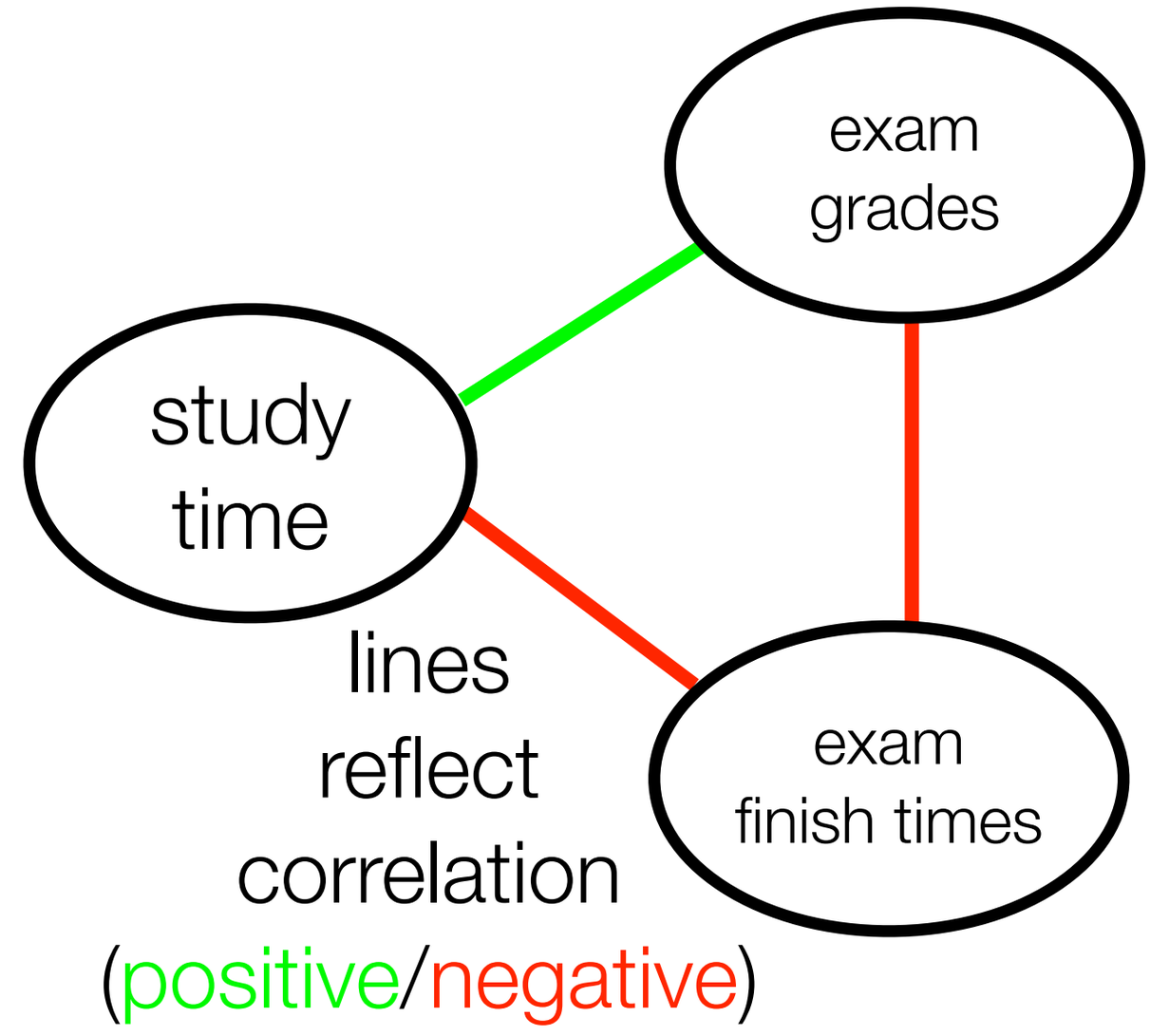
Correlations may reflect causal relations or the effects of common causes



1 *How does correlation relate to causation?*

Correlation and causation

- Correlations can sometimes imply the wrong causal relation
- Negative correlation between exam grades and exam finishing time
  - Might be interpreted to mean that finishing the exam faster will improve grades!
- So if we *only* measured exam grades & finish times in this study, we might fool ourselves!



1 *How does correlation relate to causation?*

Natural experiments

- Ideally, if we want to be able to draw stronger inferences about causal relationships between variables, we would run a **randomized controlled experiment**.
- But this isn't always possible! (e.g., randomly assigning smoking habits to individuals and tracking them over a long time)
- Instead, a still-powerful alternative is a **natural experiment.**
  - A natural experiment is an empirical study in which individuals (or clusters of individuals) are exposed to the experimental and control conditions that are determined by nature or by other factors outside the control of the investigators.

1 *How does correlation relate to causation?*

Example: Natural experiment exposing relationship between smoking and heart disease

- In Helena, Montana a smoking ban was in effect in all public spaces, including bars and restaurants, during the six-month period from June 2002 to December 2002.
- Helena is geographically isolated and served by only one hospital. The investigators observed that the rate of heart attacks dropped by 40% while the smoking ban was in effect.
- Opponents of the law prevailed in getting the enforcement of the law suspended after six months, after which the rate of heart attacks went back up.
- This study was an example of a natural experiment, called a case-crossover experiment, where the exposure is removed for a time and then returned.
- However, the inability to control variables in natural experiments can impede investigators from drawing firm conclusions. Critics argued that the particularly large percentage fluctuation in the rate of myocardial infarction was likely due to chance, given the small population size.

Today

Lecture 20: broader methodological trends in psychology

1 → 2 → 3

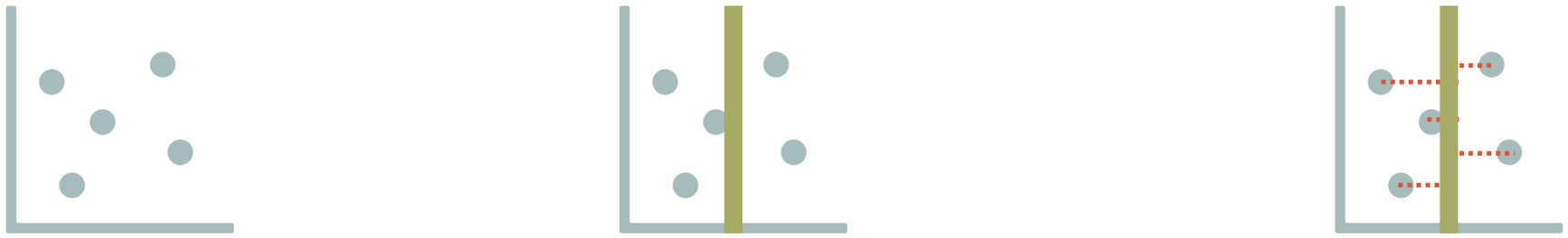
How does correlation relate to causation?

Problems with status quo: p-hacking, HARK-ing

Open science and computational reproducibility

2

Problems with the status quo: p-hacking, HARK-ing



data = model + error

what we actually observe

what we expect to observe

difference between expected and observed

What is the General Linear Model (GLM)?

A general linear model is a specific type of statistical model in which the values of a dependent/outcome variable is determined by a linear combination of independent predictor variables that are each multiplied by a weight (often represented by the letter ***b*** or Greek letter "beta," β).

$$Y_i = b_0 + b_1 X_i + e_i$$

observed value of outcome variable
e.g., thumb length

intercept

slope

value of explanatory variable
e.g., height

error

$\hat{Y}_i$ ***predicted*** *value of outcome variable*

Y_i ***observed*** *value of outcome variable*

Have you heard about a "replication crisis" in psychology?

Poll Everywhere

2 *Problems with the status quo: p-hacking, HARK-ing*

The Atlantic — Popular · Latest · Sections · Magazine · More

SCIENCE

Psychology's Replication Crisis Is Running Out of Excuses

Another big project has found that only half of studies can be repeated. And this time, the usual explanations fall flat.

ED YONG NOV 19, 2018

nature — international weekly journal of science

News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video

Volume 533 | Issue 7604 | News Feature | Article

TURE | NEWS FEATURE

,500 scientists lift the lid on reproducibility

rvey sheds light on the 'crisis' rocking research.

nya Baker

May 2016 | Corrected: 28 July 2016

PDF | Rights & Permissions

A reproducibility crisis?

The headlines were hard to miss: Psychology, they proclaimed, is in
crisis.

The New York Times

SCIENCE TIMES AT 40

Essay: The Experiments Are Fascinating. But Nobody Can Repeat The

Science is mired in a "replication" crisis. Fixing it will n

SCI
AM

ES | MIND | HEALTH | TECH | SUSTAINABILITY | EDUCATION | VIDEO | PODCASTS | BLOGS | PU

MIND

Massive International Project Raises Questions about the Validity of Psychology Research

Failures of traditional NHST to provide strong evidence against pseudoscientific discoveries (e.g., precognition, clairvoyance)

Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect

Daryl J. Bem
Cornell University

The term *psi* denotes anomalous processes of information or energy transfer that are currently unexplained in terms of known physical or biological mechanisms. Two variants of psi are *precognition* (conscious cognitive awareness) and *premonition* (affective apprehension) of a future event that could not otherwise be anticipated through any known inferential process. Precognition and premonition are themselves special cases of a more general phenomenon: the anomalous retroactive influence of some future event on an individual's current responses, whether those responses are conscious or nonconscious, cognitive or affective. This article reports 9 experiments, involving more than 1,000 participants, that test for retroactive influence by "time-reversing" well-established psychological effects so that the individual's responses are obtained before the putatively causal stimulus events occur. Data are presented for 4 time-reversed effects: precognitive approach to erotic stimuli and precognitive avoidance of negative stimuli; retroactive priming; retroactive habituation, and retroactive facilitation of recall. The mean effect size (*d*) in psi performance across all 9 experiments was 0.22, and all but one of the experiments yielded statistically significant results. The individual-difference variable of stimulus seeking, a component of extraversion, was significantly correlated with psi performance in 5 of the experiments, with participants who scored above the midpoint on a scale of stimulus seeking achieving a mean effect size of 0.13. Skepticism about psi, issues of replication, and theories of psi are also discussed.

Keywords: psi, parapsychology, ESP, precognition, retrocausation

Journal of Personality and Social Psychology (2011)

Renewed scrutiny of typical procedures used during data collection and analysis in psychology

General Article

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Joseph P. Simmons[1], Leif D. Nelson[2], and Uri Simonsohn[1]
[1]The Wharton School, University of Pennsylvania, and [2]Haas School of Business, University of California, Berkeley

Table 1. Likelihood of Obtaining a False-Positive Result

	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

Psychological Science (2011)

Cognitive biases in statistical/scientific reasoning

- "The first principle is that you must not fool yourself and you are the easiest person to fool"- R. Feynman
- We pay more attention to information that confirms our hypotheses or biases versus those that disconfirm them.
- We are more likely to overlook errors that confirm our pre-existing ideas.
- We fail to consider alternative hypotheses that could explain the data.

2 *Problems with the status quo: p-hacking, HARK-ing*

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

PLoS Medicine August 2005 | Volume 2 | Issue 8 | e124

John Ioannidis

"There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. ... Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. "

Statistical power (our ability to detect real effects) remains low in many areas of social science

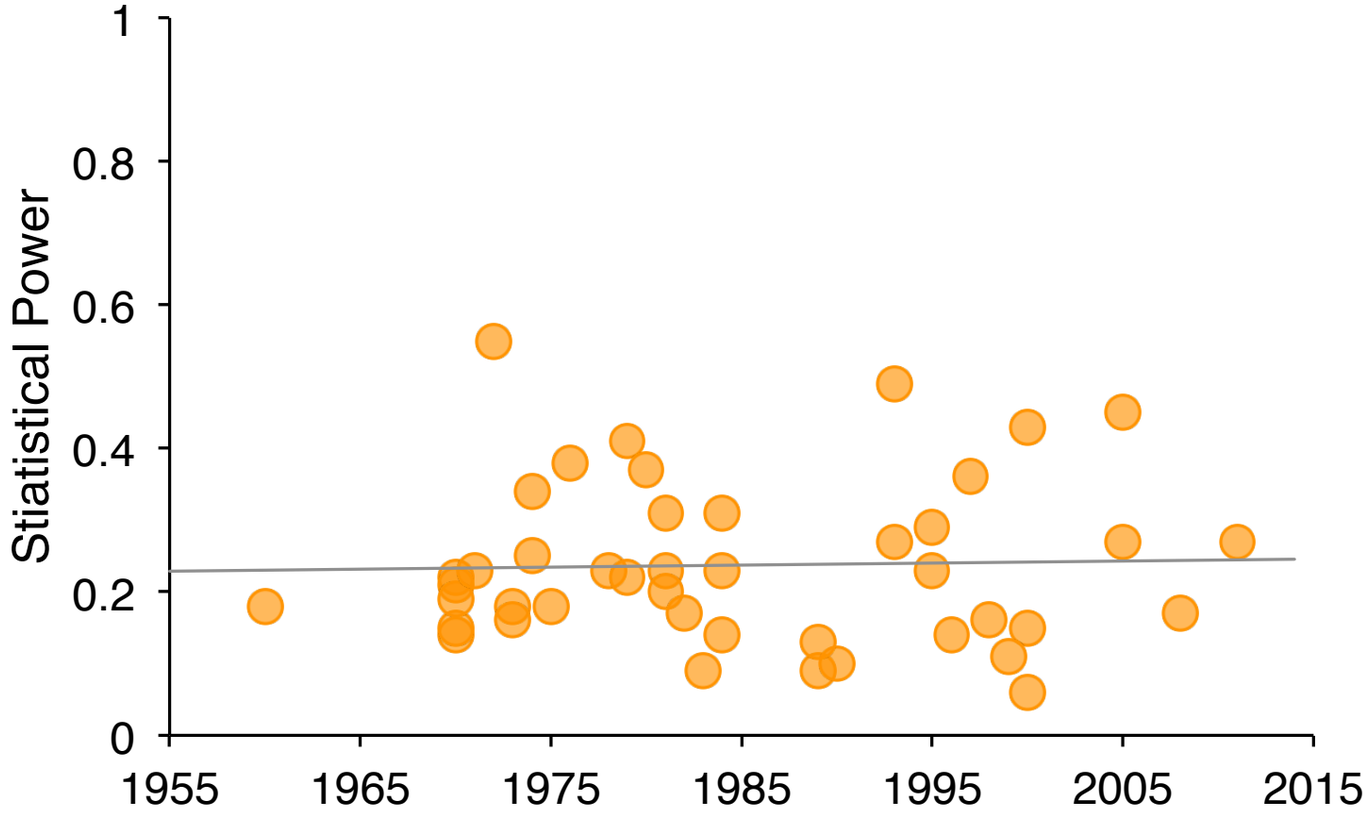


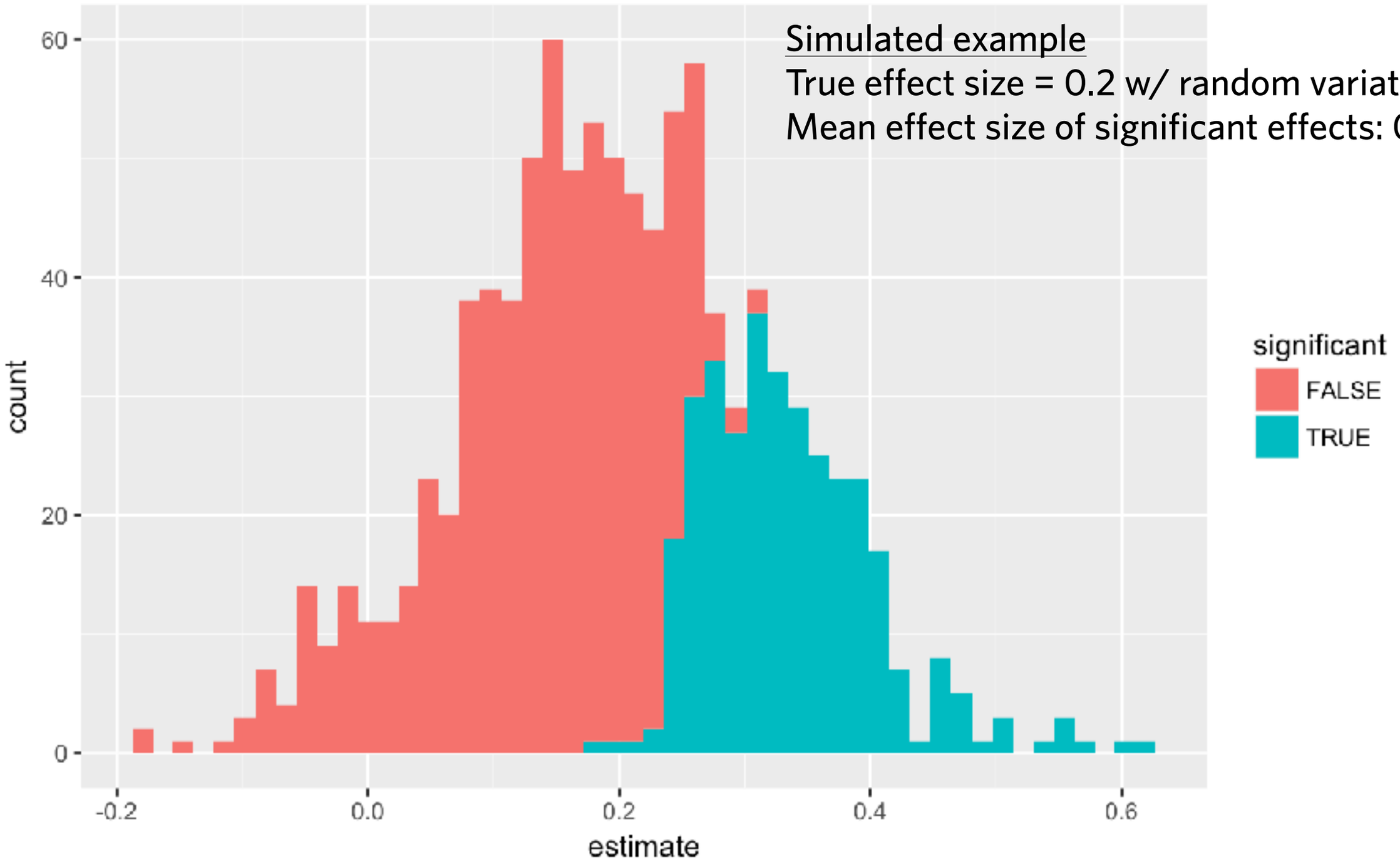
FIGURE 1. Average statistical power from 44 reviews of papers published in journals in the social and behavioral sciences between 1960 and 2011. Data are power to detect small effect sizes ($d = 0.2$), assuming a false positive rate of $\alpha = 0.05$, and indicate both very low power (mean = 0.24) but also no increase over time ($R^2 = 0.00097$).

Smaldino & McElreath, 2016, https://arxiv.org/abs/1605.09511

The winner's curse: how the size of estimated effects is inflated by Null Hypothesis Significance Testing

- In economics, for certain types of auctions (where the value is the same for everyone, like a jar of quarters, and the bids are private), the winner almost always pays more than the good is worth.
- In statistics, the effect size estimated from significant results (i.e. the winners) is almost always an overestimate of the true effect size.

The winner's curse: how the size of estimated effects is inflated by NHST

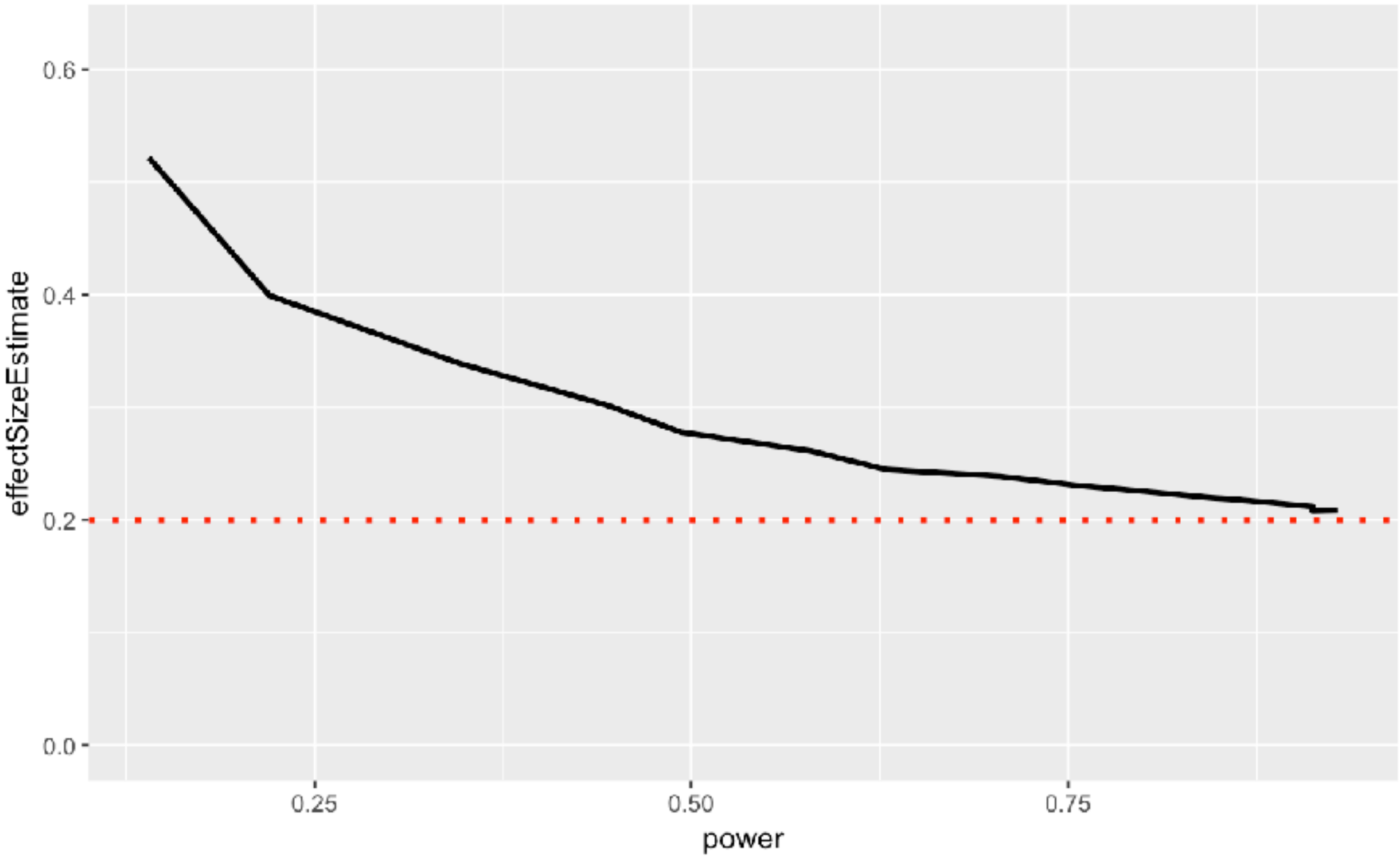


Simulated example
True effect size = 0.2 w/ random variation
Mean effect size of significant effects: 0.33

2

Problems with the status quo: p-hacking, HARK-ing

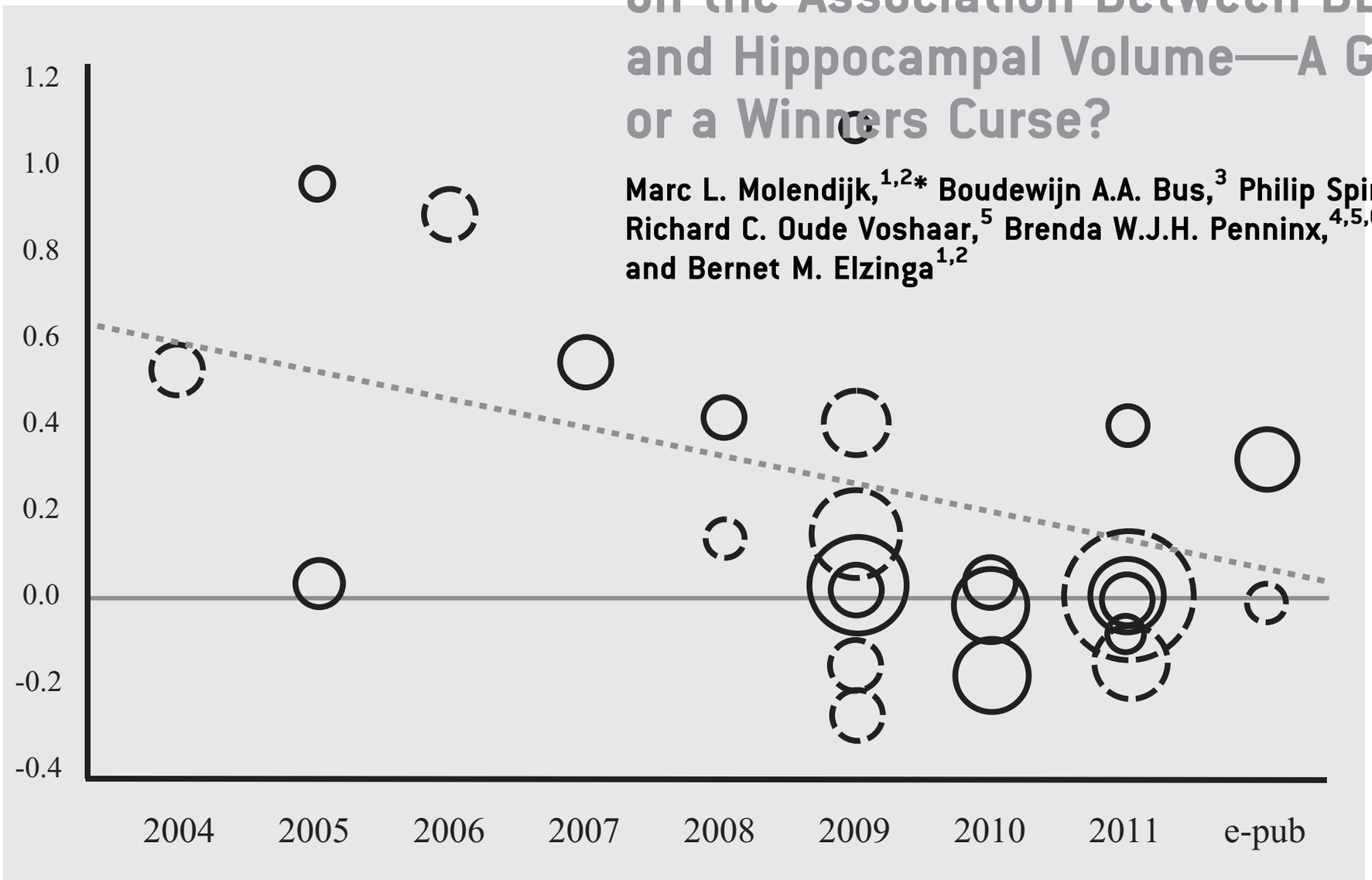
How much NHST inflates effect size estimates lessens as the power of individual studies increases.



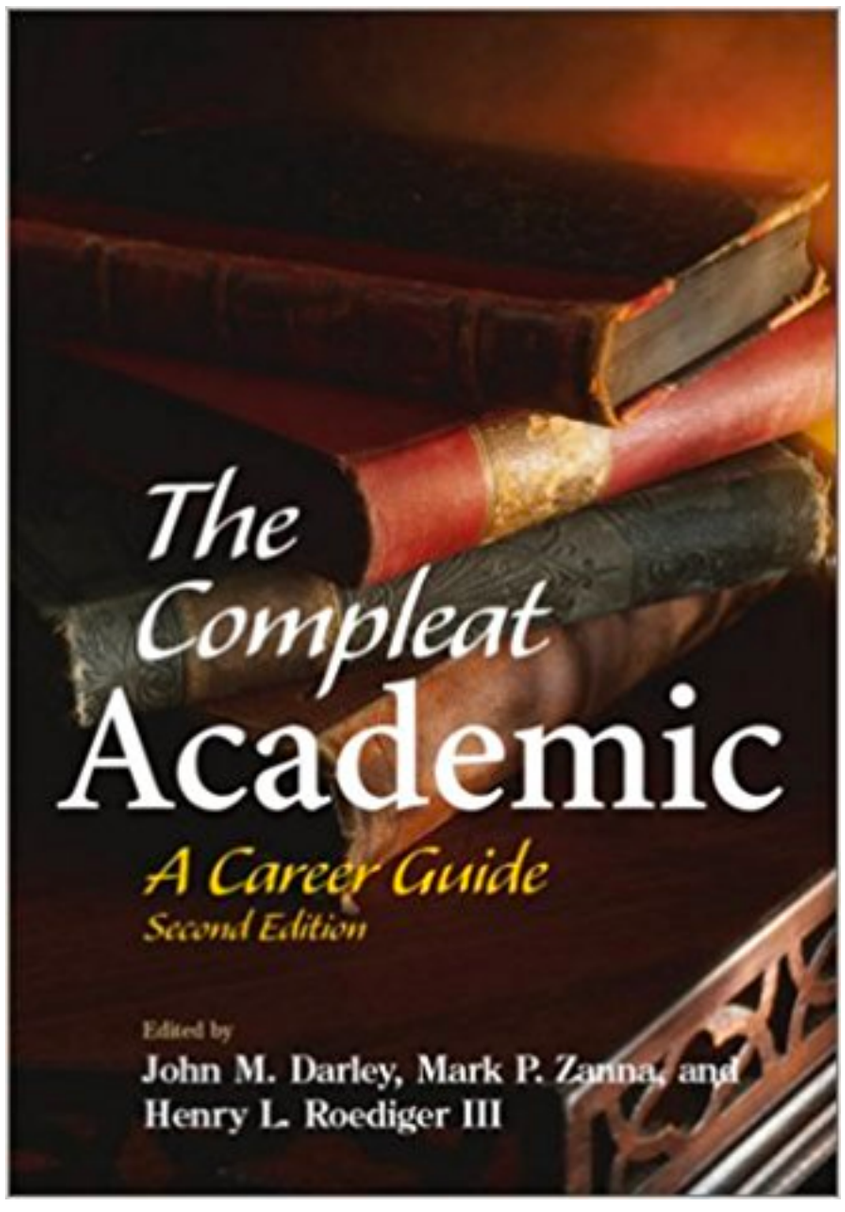
How much NHST inflates effect size estimates lessens as the power of individual studies increases.

A Systematic Review and Meta-Analysis on the Association Between BDNF val^{66}met and Hippocampal Volume—A Genuine Effect or a Winners Curse?

Marc L. Molendijk,[1,2]* Boudewijn A.A. Bus,[3] Philip Spinhoven,[1,2,4] Anna Kaimatzoglou,[1] Richard C. Oude Voshaar,[5] Brenda W.J.H. Penninx,[4,5,6] Marinus H. van IJzendoorn,[7,8] and Bernet M. Elzinga[1,2]



What is the status quo? Advice for early-career psychologists.



A new career in academia can be a challenge. While academia's formal rules are published in faculty handbooks, its implicit rules are often difficult to discern. Like the first edition, this new and expanded volume of The Compleat Academic is filled with practical and valuable advice to help new academics set the best course for a lasting and vibrant career.

https://www.apa.org/pubs/books/4316014.aspx

2

Problems with the status quo: p-hacking, HARK-ing

Career advice from Daryl J. Bem:

"HARKing"

Which Article Should You Write?
There are two possible articles you can write: (a) the article you planned to write when you designed your study or (b) the article that makes the most sense now that you have seen the results. They are rarely the same, and the correct answer is (b).

"p-hacking"

re Data Analysis: Examine them from every angle. Analyze the sexes separately. Make up new composite indexes. If a datum suggests a new hypothesis, try to find additional evidence for it elsewhere in the data. If you see dim traces of interesting patterns, try to reorganize the data to bring them into bolder relief. If there are participants you don't like, or trials, observers, or interviewers who gave you anomalous results, drop them (temporarily). Go on a fishing expedition for something— anything —interesting.

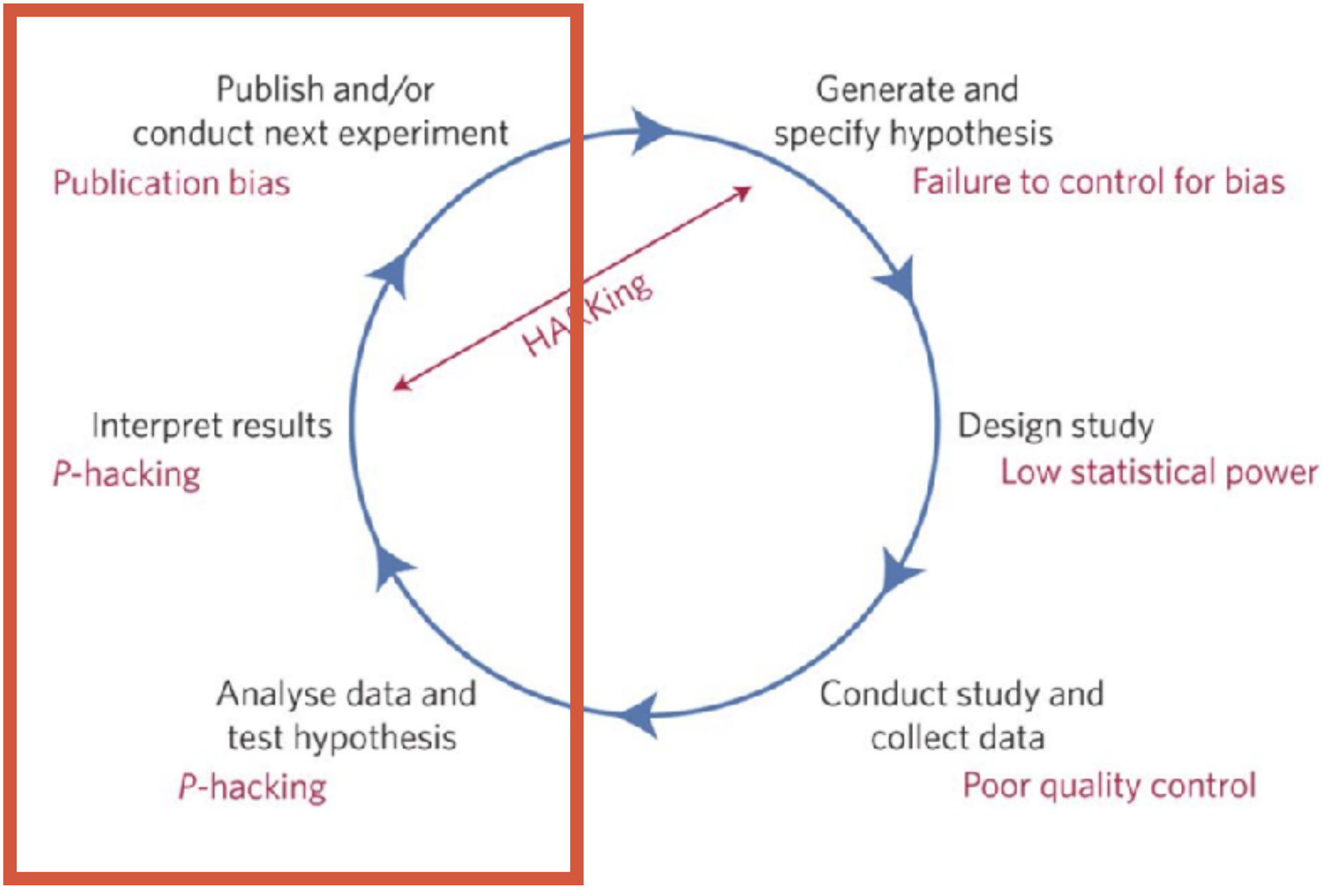
https://www.apa.org/pubs/books/4316214.aspx http://neuroanatomy.wp.ox.ac.uk/2017/11/oxford-reproducibility-lectures-dorothy-bishop/

"HARKing"

- "Hypothesizing After the Results are Known" (Kerr, 1988)
- Why is this a problem?
  - It can turn false positives into "theory."
    - A post-hoc conclusion gets re-framed as an *A PRIORI* hypothesis
    - A theory that is re-written to fit the facts is not a very powerful theory, and not expected to generalize well to new observations.
  - It becomes impossible to disconfirm bad ideas

2 *Problems with the status quo: p-hacking, HARK-ing*

The idealized scientific process and how common practices undermine it.



where statistical practices intervene

This is why we asked you all to pre-register your research question and analysis approach *before* seeing the data for your final projects.

2 *Problems with the status quo: p-hacking, HARK-ing*

Play the "p-hacking game:" find a significant effect & win!

- Go to: https://projects.fivethirtyeight.com/p-hacking/
- Last names starting with letters A-L:
  - Find evidence that the U.S. economy is better when Republicans are in office.
- Last names starting with letters M-Z:
  - Find evidence that the U.S. economy is better when Democrats are in office.
- Tell us on PollEverywhere once you have a significant effect!

Poll Everywhere

Play the "p-hacking game:" find a significant effect and win! Tell us once you have a significant effect and how you got it!

Why don't we believe in ESP (yet)?

- Bem's paper shows evidence of p-hacking
  - Sample sizes varied across studies
  - Different studies appear to have been lumped together or split apart
  - The studies allow many different hypotheses, and it's not clear which were planned in advance
  - Uses one-tailed tests even when it's not clear that there was a directional prediction (so alpha is really 0.1)
  - Most of the p-values are very close to 0.5
  - It's not clear how many other studies were run but not reported

http://www.talyarkoni.org/blog/2011/01/10/the-psychology-of-parapsychology-or-why-good-researchers-publishing-good-articles-in-good-journals-can-still-get-it-totally-wrong/

Today

Lecture 19: broader methodological trends in psychology

How does correlation relate to causation?

Problems with status quo: p-hacking, HARK-ing

Open science and computational reproducibility

3 *Open science and computational reproducibility*

How do we overcome these problems with the status quo?

- Don't p-hack
- Pre-register your research questions & analysis plan
- Publish both positive and negative results (do not make publication dependent on the results)
- Replicate studies whenever possible
- Make your code and data available so your future self & other researchers can verify and build upon your work
  - *You've been learning the foundations of "computational reproducibility" throughout the quarter by writing your analysis workflows in R!*

http://www.talyarkoni.org/blog/2011/01/10/the-psychology-of-parapsychology-or-why-good-researchers-publishing-good-articles-in-good-journals-can-still-get-it-totally-wrong/

What does reproducibility mean?

		DATA	
		Same	**Different**
CODE	**Same**	**Reproducible**	**Replicable**
	Different	**Robust**	**Generalizable**

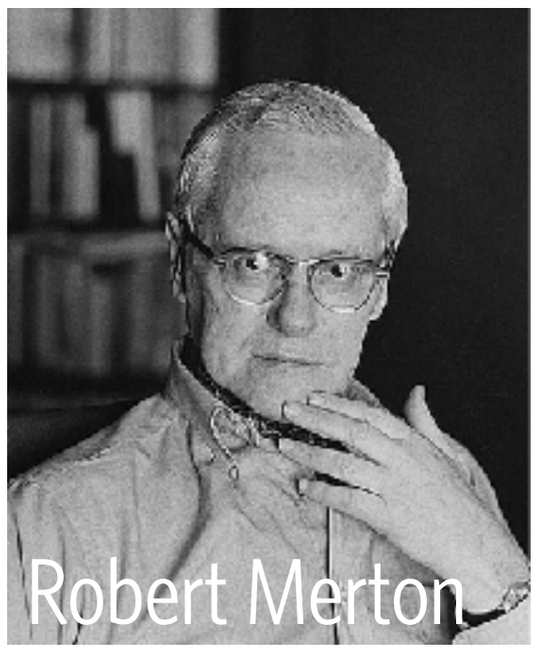
https://figshare.com/articles/Publishing_a_reproducible_paper/5440621

What is "open science?"

Open science refers to the set of beliefs, research practices, and policies organized around the central roles of transparency and verifiability in scientific practice.

("Take nobody's word for it")

Mertonian scientific norms



- **Communism**: all scientists own all science
- **Universalism**: scientific validity is independent of individuals or institutions
- **Disinterestedness**: scientific gains are for the common good, not the good of the individuals
- **Skepticism**: claims should be subject to organized scrutiny prior to being adopted
- **These values are fundamentally aligned with openness!**

Preregistration

- Pre-registration entails writing up the goals of study, testable hypotheses, analysis approach ahead of actually working with data & seeing the results.
- Helps to protect self against inadvertent p-hacking and HARKing.
- More detailed plans help reduce "researcher degrees of freedom."
- Distinguishes between "exploratory" and "confirmatory" research.
  - Exploratory research is data-driven.
  - Confirmatory research is theory-driven.

What does reproducibility mean?

		DATA	
		Same	**Different**
CODE	**Same**	**Reproducible**	**Replicable**
	Different	**Robust**	**Generalizable**

https://figshare.com/articles/Publishing_a_reproducible_paper/5440621

Foundations of computational reproducibility

- Use and share data formats that are easy to open across different computing environments
  - e.g., CSV files that can be imported as tidy data frames.
- Write your data analysis pipeline using a scripting language that can be executed in the same way by any computer every time (rather than using a point-and-click interface)
  - e.g., RMarkdown Notebooks
- **In this class, you have learned the foundational tools for computationally reproducible statistical analyses!**

3

Open science and computational reproducibility

Science is an *iterative process*, not a collection of facts.

Statistics gives us tools for turning our research questions and hypotheses into testable predictions, and for evaluating the strength of the evidence we have.

A single study generally doesn't close the door on a research question, but rather open up ***new questions for future research.***

🙌 **In carrying out your final project, you are participating in this process! Congrats!**

Recap

Lecture 20: broader methodological trends in psychology

1 → 2 → 3

How does correlation relate to causation?

"Correlation does not imply causation, but it's a pretty good hint!"

Problems with status quo: p-hacking, HARK-ing

Inflate the actual false-positive rate above the presumed one of 0.05.

Open science and computational reproducibility

Share your data in tidy format & use scriptable analysis workflows to ensure computational reproducibility!


Thank You
:)

🙏