Pseudoscience and the Replication Crisis in Psychology

Cognition Lab 04/17/2020

What is Science?

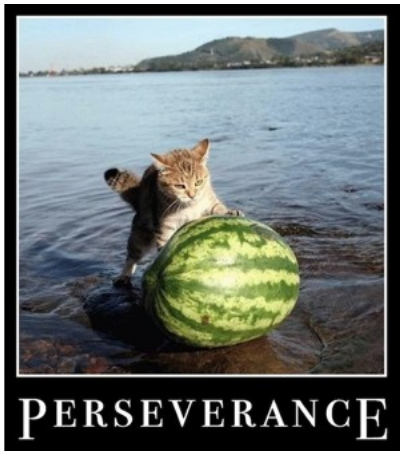
Science is the systematic study of natural events and conditions.

In general, all scientific subjects can be broken down into three areas—**life science, Earth science, and physical science**.

What is Psychology?

Psychology:
the science of human & non-human animal behavior





Goal: gather an organized body of knowledge

Psychological research seeks scientific explanation for behavior

RUTGERS

What is Science?

Scientific evidence must be *open for all to observe*. It must be measurable and not just one person's opinion or guess.

Empirical evidence includes *observations and measurements*. It includes the data people gather and test to evaluate scientific explanations.

The **scientist** may use *specialized tools* to make observations and take measurements.

What is Pseudoscience?

Something that looks like science and pretends to be science … but isn't!

- Proto-science
- Speculative science
- **Pseudoscience!** It relies on …
  1. Scientific language
  2. Appeal to authority
  3. Forceful arguments
  4. Appeal to intuition
  5. Popular media

Examples

RUTGERS

Pseudoscience Related to **Space and Astronomy**

- 2012 Phenomenon - end of the world
- Crop Circles - made by aliens
- Face on Mars - A rock formation: evidence of life

Pseudoscience Related to the **Earth**

- Cryptozoology - search for Bigfoot (Yeti), Loch Ness monster, El Chupacabra

Pseudoscience Related to the **Paranormal**

- Channeling - Involves communicating with a spirit through a person
- Electronic Voice Phenomenon (EVP) - Used by ghost hunters to record messages from spirits
- Extra-sensory Perception (ESP) - clairvoyance, telepathy, levitation
- Numerology - A belief in a relationship between certain numbers and people or events

Pseudoscience Related to **Psychology**

- Conversion Therapy - change sexual preference (homosexual to heterosexual)
- Hypnosis - method of deep relaxation where the subject is open to suggestions
- Psychoanalysis – Freud

Other Pseudoscience Examples

- Feng Shui - A Chinese method of arranging objects and living spaces in a certain way
- Handwriting Analysis - A belief that a person's personality can be seen in their handwriting

RUTGERS

Science & Pseudoscience – Popper's Demarcation Criterion

Karl Popper

"The Logic of Scientific Discovery" (1934, 1959)



Popper's Demarcation Criterion

Popper raised the *demarcation problem*:

When should a theory be ranked as scientific?

Is there a **criterion** for the scientific status of a theory?

Scientific / Pseudoscientific

Interpretation Vs Explanation

The **“explanatory power”** of a pseudoscientific theory is just a disguise of its **“interpreting power”**.
(Can this theory *predict* behavior?)

This **apparent strength** - capable of explaining everything - is in fact a **weakness**.

Popper's Demarcation Criterion: Falsifiability

Consequently, Popper proposed

The criterion of the scientific status of a theory is its

falsifiability,

or *refutability*,

or *testability*.

Science should follow a methodology based on falsifiability

We can never prove a theory true…

… we should *be able* to prove it false!

Is this criterion always adequate?

Criteria of Adequacy

Falsifiability

Scope

Fruitfulness

Conservatism

Simplicity

What's their significance?

Falsifiability

- Necessary condition for being scientific
- Possible candidate for knowledge
- Minimal condition for further study

Scope, Fruitfulness, Conservatism, Simplicity

- Involved in systematizing, unifying & developing scientific knowledge.

Pseudoscience and Falsifiability

Pseudoscience theories refuse falsifiability
(beware of certainty…)
“Is there any way we could prove this to be false?”
”No” --> Pseudoscience
As Popper said: “not only these pseudoscience theories are not right… *they are not even wrong*!”

Pseudoscience and Falsifiability: Autism's therapy

"Is there any way we could prove this to be false?"
"No" --> Pseudoscience
"Is there any way this therapy could be *not the best one for the case at hand*?"
"No" --> Pseudoscience

Neurotribes by Steve Silberman

"On the Rosas' next pilgrimage to Los Altos, the doctor inevitably brought up chelation. But this time Craig challenged him. "Wait a second," he told the doctor. "You're telling me that the recommended course of action for a low reading of mercury toxicity is chelation?" "Yes," the doctor replied. "And the recommended course of action for high mercury toxicity is chelation?" The doctor nodded yes again. **Finally Craig asked him, "Is there any sort of outcome that would contraindicate chelation?" And the doctor said, "No."** At that point, Craig and Shannon said, "Thank you very much," walked out of the doctor's office, and never went back to Los Altos to see him again."

Pseudoscience and Psychology

We saw several fields that could be classified as pseudoscience

Some of them are related to psychology:

- Conversion therapy
- Hypnosys

Bem (2011) tried to present “pre-cognition” as a scientific phenomenon:

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

Pseudoscience and Psychology

Hypnosis: is there a way to test it?

When hypnosis fails: some people are not good subjects for hypnosis…

Pre-cognition: Bem tried to test it scientifically

The problem is that he “succeeded”

Pseudoscience and Psychology

Researchers asked: how is that possible?



By Daniel Engber

TWEET SHARE

Daryl Bem Proved ESP Is Real

Which means science is broken.

MAY 17, 2017 • COVER STORY

2011: the crisis of Psychology

- Daryl Bem's study on precognition
- Diederik Stapel
- "False-Positive Psychology" (Simmons et al. 2011), → "listening to a song can change a person's age"
- Doyen et al. (2012) **failure to replicate** one of the most famous findings in social psychology: priming people with elderly stereotypes made them walk more slowly (Bargh et al. 1996)

The "Replication Crisis"

If our scientific methods allow us to:

- Prove phenomena we know to be wrong

- Cheat easily

- Prove anything we want

There must be something wrong with them

An incorrect methodology blurs the line between science in pseudoscience

Significance?

What does it mean for a result to be significant?

Researcher calculate a statistics called “p-value”

“How likely is it that this result could have happened by chance”

If $p < .05$ → result is significant

Still, there are 5 in 100 chances (1 in 20) that you may be wrong

Significance?

There is one in twenty chance that you may be wrong

Imagine you have phenomenon happening by chance…

You test it 100 times…

On average, 5 times out of 100 it will be significant

Significance?

Do you need to run a study 100 times to have 100 tests?
Or 20 times to have 20 tests?

No

P-hacking:

- You can measure many variables at the same time
- And then cross-test all of them! Now you have increased the number of tests with just one study
- Some of these test will be significant..
- You report only those that are significant

Significance?

Do you need to run a study 100 times? Or 20 times?

No

P-values are sensitive to sample size

The bigger your sample, the higher the chance your p-value will be small

Problems identified

Paradox:

Most published studies are significant…

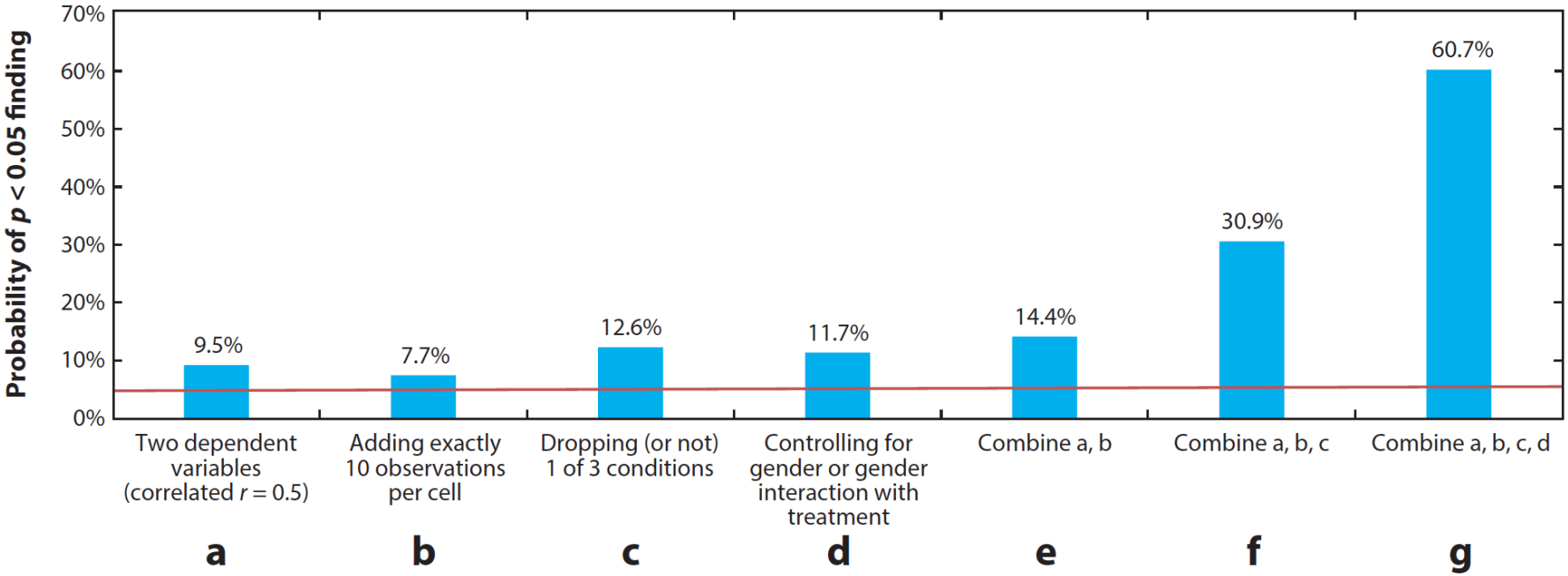
Most published studies are underpowered

Is there an explanation to this apparent paradox?

Solution to the paradox: p-hacking

RUTGERS

How to push a significant result



Controversies and personal attacks

RUTGERS

Brian Wansink – Cornell



Amy Cuddy - Harvard



Discovering p-hacking in the past

RUTGERS

Replication

Open Science Collaboration (Open Sci. Collab. 2015) replicate the procedures from 100 published psychology studies

They replicated about 60%



Preventing p-hacking in the future

RUTGERS

Disclosure: explain exactly all the analysis you did

Pre-registration









Anything else

Are p-value really bad?

The problem is that p-values are

- Abused
- Misinterpreted

Solution: Promote statistical education

Results section

More on presenting Figures/Tables

RUTGERS

Figures/Tables often help explain things in fewer words

The reader should not have to refer to the text to understand your tables and figures: **there should be explanatory figure/table captions**

Figures and tables are labeled independently

- E.g. Figure 1 then later Table 1 (not Figure 1 then later Table 2!)

Axes on graphs should be clearly labeled

Units of measurement should be clear



Don't refer to 'graphs', 'diagrams', 'pictures', photographs' etc. in your main report text; use the terms '**Figure' and/or 'Table**'

Results

This section provides the reader with a clear, concise summary of the data you collected and the results of any statistical tests

Normally two parts to the results (not sub-sections though!)

1. Describe/Present the descriptive statistics
   - Means, standard deviations etc. for each group/condition
2. Follow with inferential statistical tests you have performed to examine the data
   - t-tests, ANOVA results, correlations etc.
   - E.g. 'An independent samples *t*-test was performed to examine any differences between the male and female groups. Males performed significantly faster than females ($t(29) = 3.65$, $p < .05$)
   - *Note that italics are used for a statistic and for probability*

Descriptive Statistics

When referring to descriptive/summary statistics try to be consistent with APA style

- *M* represents a mean. *Mdn* represents a median
- *SD* represents the standard deviation. *SE* represents standard error.
  - Remember to use UNITS
  - e.g. 'Males were faster than females on the image rotation task (Males: *M* = 20.1, *SD* = 3.6 msecs; Females: *M* = 30.5, *SD* = 7.3 msecs)

A graphical representation of summary statistics is best when there are a few means to compare

If there are lots of means/summary statistics then you may want to present them in table form

Results- Reporting Inferential Statistics /probabilites

RUTGERS

Always report the following information when reporting a statistic analysis

- the value of the test statistic
- degrees of freedom (or number of subjects minus one)
- one- or two- tailed test
- the observed *p*-value
- whether the test result was significant
  - Exact probability: E.g. 'The mean reaction time for the younger children was significantly slower than for the older children ($t(14) = 12.56$; $p = .03$, two-tailed test)
  - Or: 'The mean reaction time for the younger children was significantly slower than for the older children ($t(14) = 12.56$; $p < .05$, two-tailed test)
- APA suggests rounding to 2 decimal places and many authors do not use a "leading zero" for probabilities
  e.g. *p= .54* rather than *p=0.542*

Results

Remember:



- Descriptive statistics should be reported in sentence form or in a table (not a list)



- Statistical tests should be reported in the standard format for that test.



- Explain what these tests mean in relation to your experiment and the predictions

Results – What *not* to do

- Don't cut and paste tables from excel into your results section
- Don't include statistical calculations (by hand or computer)
- Don't include raw data in the results section (You can put it in the appendix if necessary)
- Don't try to complicate your results- this section is often best to be left as clear and straightforward as possible
- Don't *interpret* your results at this stage – save that for the Discussion. But you *should* state your results in an understandable form. e.g.:
  - 'Males were faster than females' is a statement about your results
  - 'Males were faster than females which supports the findings of Calder (1989), but is in opposition to Broad's (1990) theory which posited that the female brain was more developed for knitting skills' is a discussion & doesn't belong in the results section.

Final Tips

A good book

- Andy Field & Graham Hole (2003). How to design and report experiments. Sage: London.
  - Chapter 9, p. 287 onwards

Write in the past-tense throughout (apart from any predictions made in the intro can be phrased in the future tense)

Avoid personal grammar ('I', 'We' etc.)

APA style

Section must be formatted in APA* style

- (*American Psychological Association)

"*Communication In Psychology*"

- At the end of PDF, example of **research paper** (like the one we are writing) in APA style, with comments
- excellent guide, maybe a bit outdated regarding APA requirements

"*Purdue guide*"

- This is an example or a **review paper**
- (up to date regarding APA– but be mindful this is a **review paper** – general structure may differ a bit from the model we are using)

"*Useful Phrases*"

- Excellent phrasebook