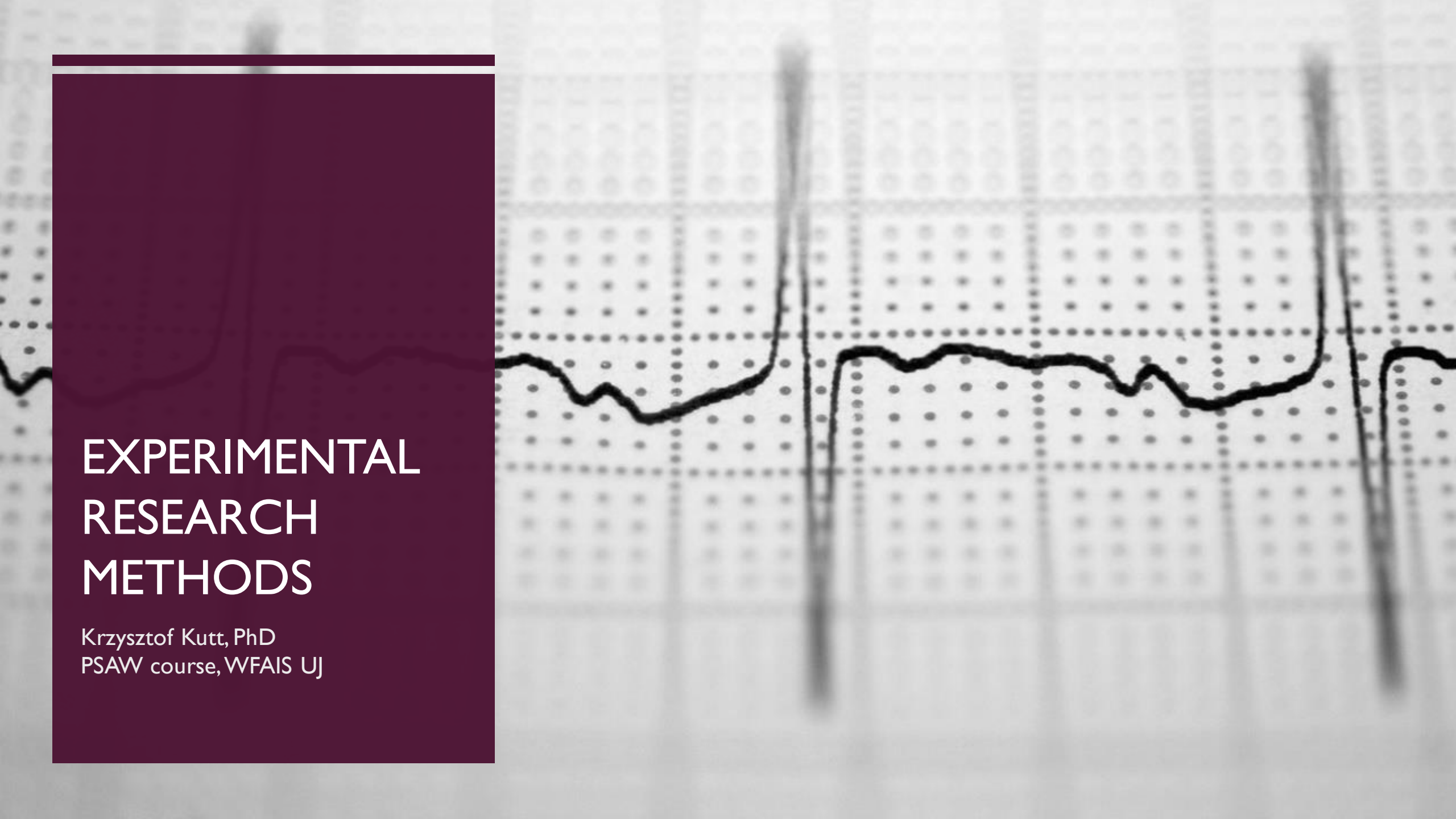


EXPERIMENTAL RESEARCH METHODS

Krzysztof Kutt, PhD
PSAW course, WFAIS UJ

PROBLEM STATEMENT

WE BEGIN WITH A PROBLEM…

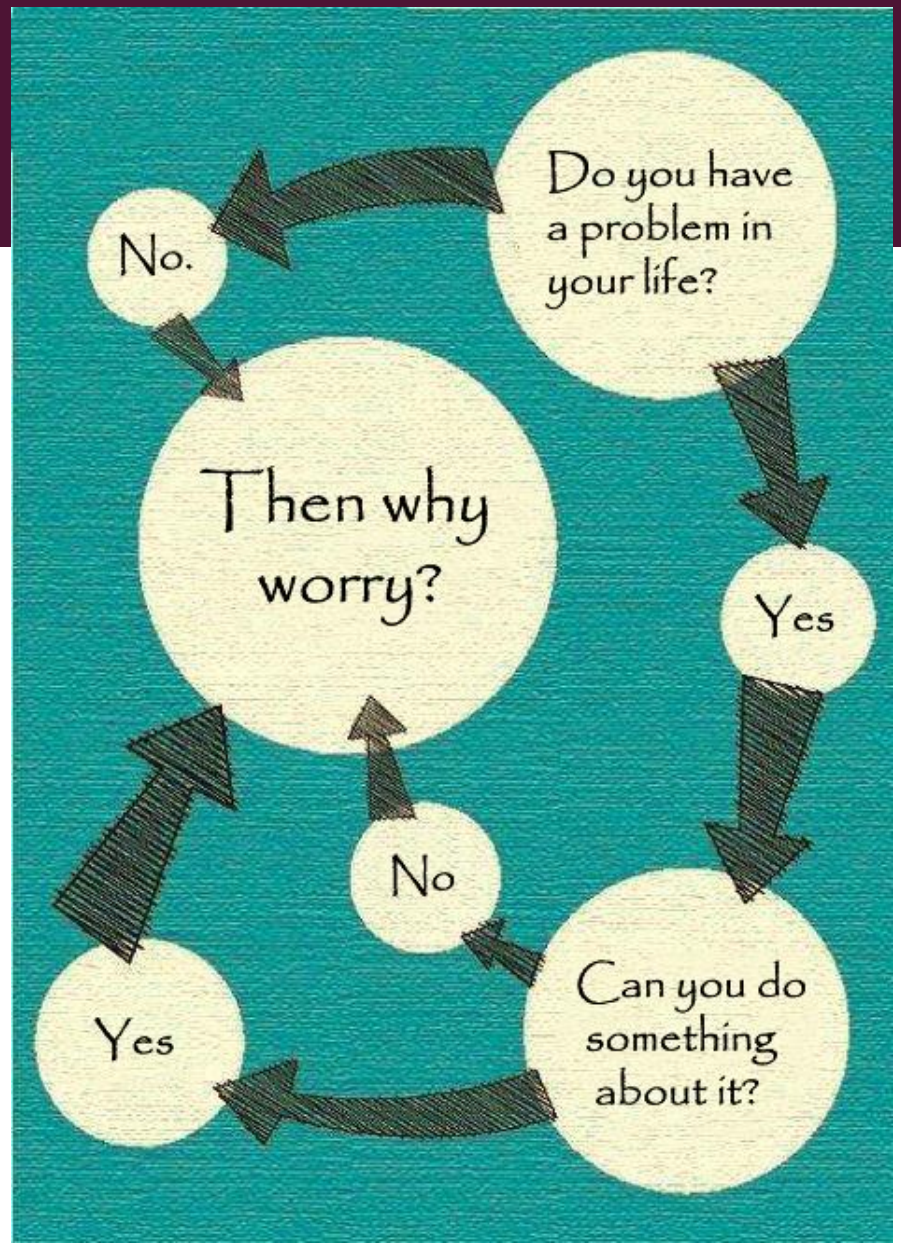
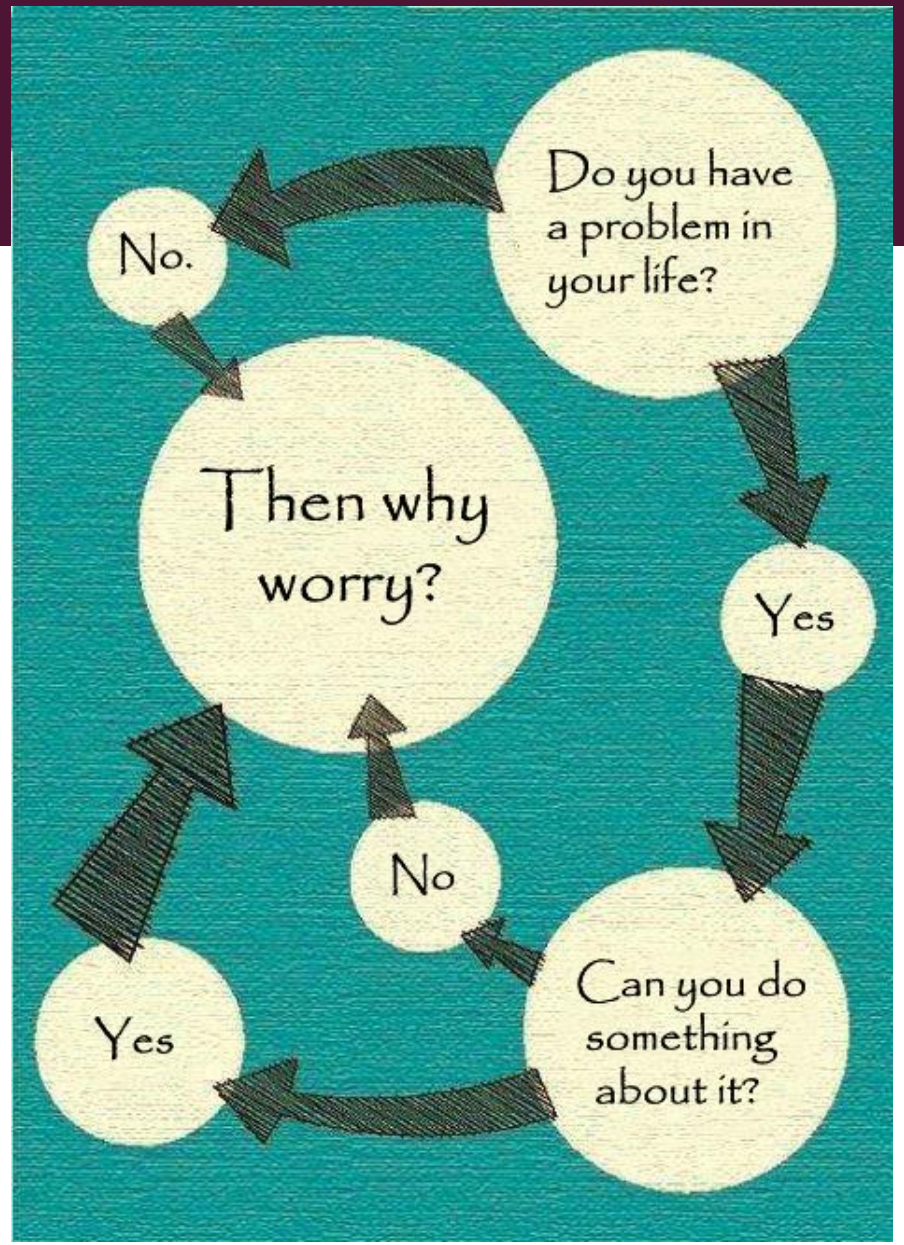


Image from wanderage.wordpress.com.

WE BEGIN WITH A PROBLEM…

Based on psychological theory:

- Replication of research already conducted,
- Extension of existing research,
- New research area (based on existing theories),
- A new research area from which a new theory will emerge.



This and following slides are based on scribbr.com.

PROBLEM STATEMENT

Aim

The overall purpose of the research:

- The aim of this study is to **determine**...
- This project aims to **explore**...
- I aim to **investigate**...

Objectives

Concrete steps:

- Qualitative methods will be used to **identify**...
- I will use surveys to **collect**...
- Using statistical analysis, the research will **measure**...

RESEARCH QUESTIONS…

Research question

Focused and specific question you will address in your research

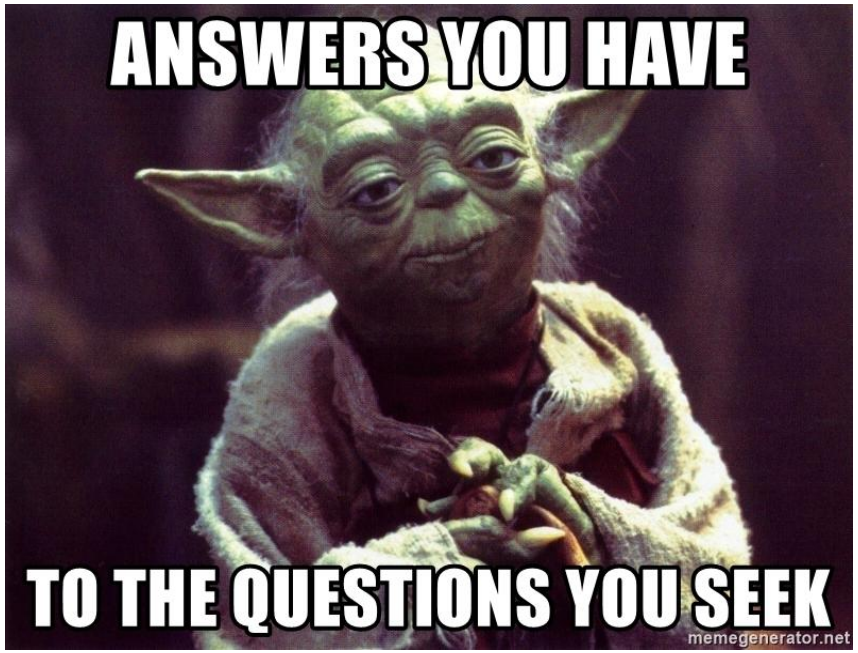
RESEARCH QUESTIONS & HYPOTHESES

Research question

Focused and specific question you will address in your research

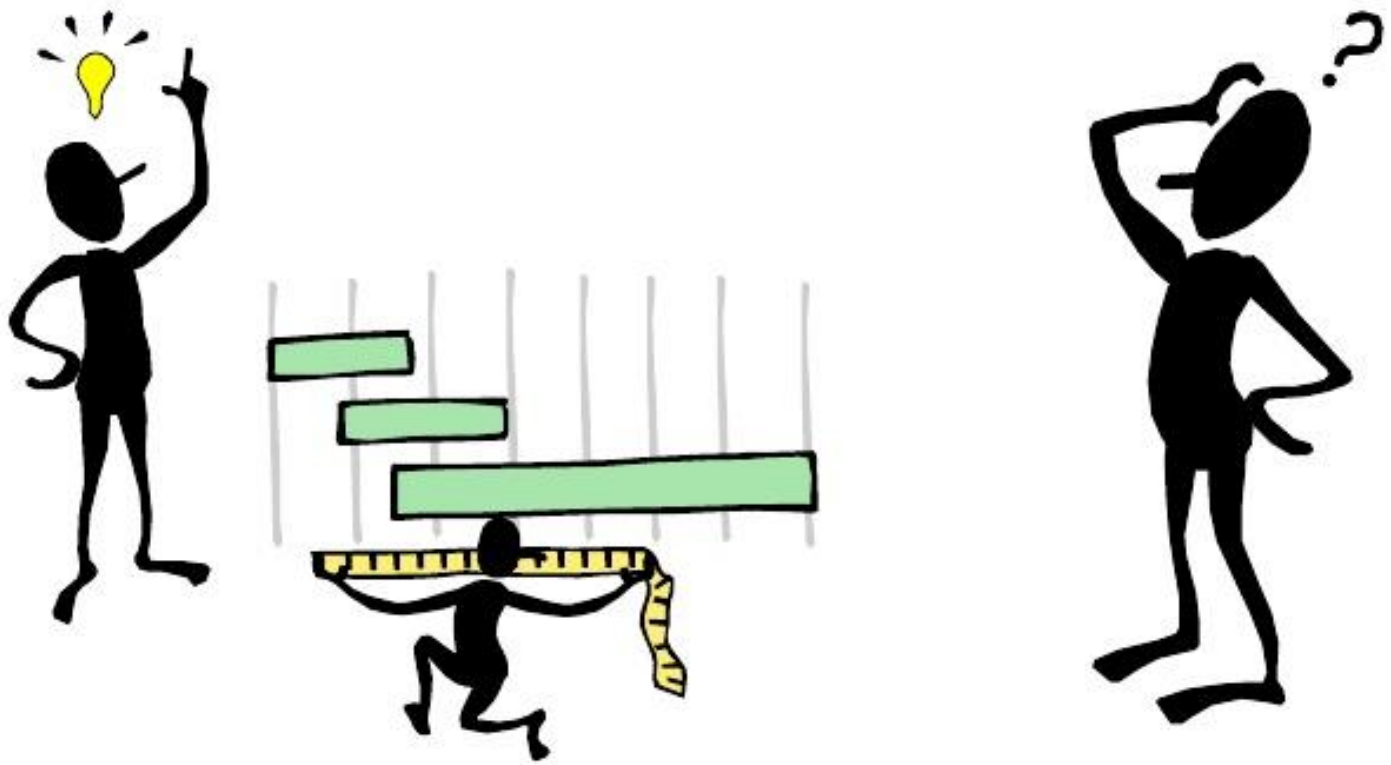
Hypothesis

A tentative answer to research question; based on existing theories, but not tested yet



VARIABLES

What is a Variable?



VARIABLES

- They values **vary**, so... they are **variables** :)
- Characteristics, properties, states of the objects we study
- What is observed and measured
- What is manipulated

VARIABLES

- They values **vary**, so... they are **variables** :)
- Characteristics, properties, states of the objects we study
- What is observed and measured
- What is manipulated

- **Quantitative** (discrete and continuous) vs **Categorical** (binary, nominal, ordinal)
- **Independent** (what is manipulated) vs **Dependent** (what is influenced)
- **Control** (e.g., light in the room was the same for all)
- **Confounding** (affect the dependent variable in an uncontrolled way)

OPERATIONALIZATION

STUDY DESIGN

CORRELATION VS EXPERIMENT?

Correlation

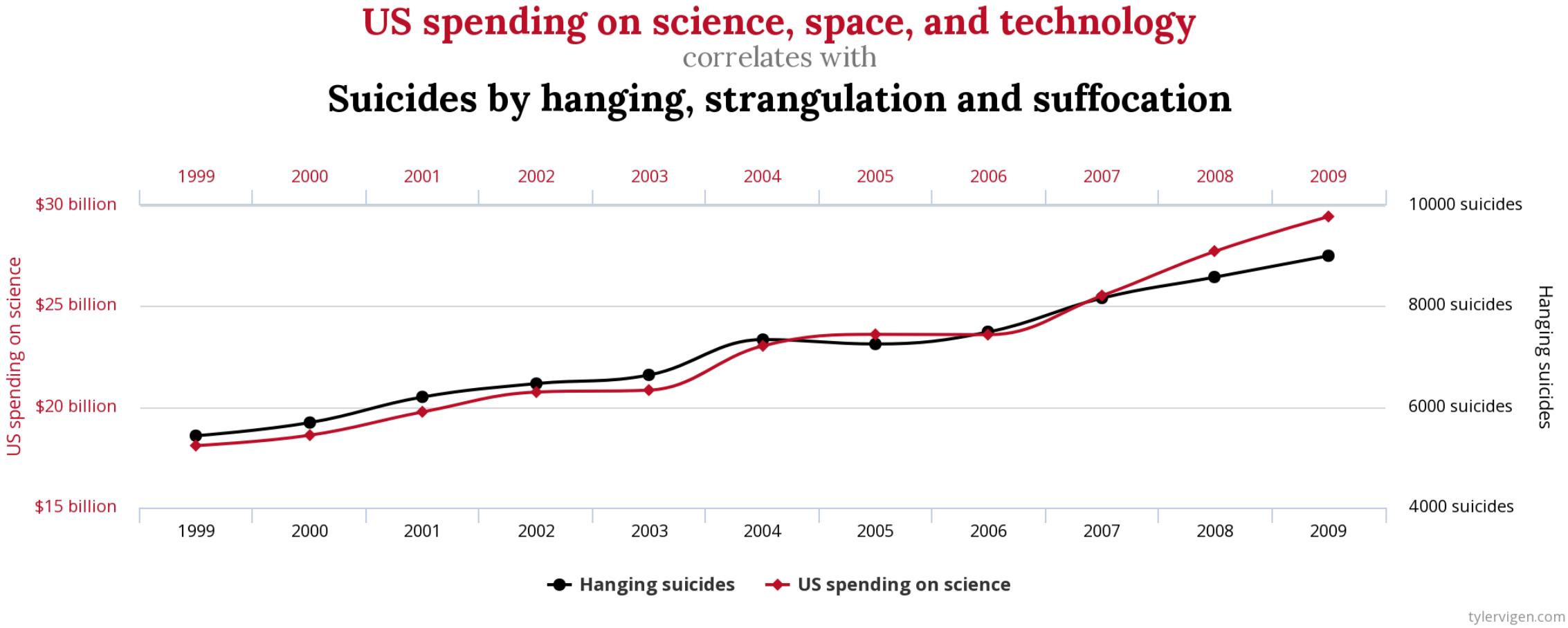
We gather data without controlling any variables

Experiment

We manipulate and control variables to determine cause and effect (we need randomization, control group, pre- and post-test; if any is missing → ***Quasi-experiment***)

CORRELATION VS EXPERIMENT?

CORRELATION VS EXPERIMENT?



See more at Spurious Correlations.

BETWEEN-SUBJECTS VS WITHIN-SUBJECTS?

Within-subjects design
The same participant tests all conditions corresponding to a variable.

Between-subjects design
Different participants are assigned to different conditions corresponding to a variable.

NNGROUP.COM NN/g

Source: nngroup.com.

HARKING

Personality and Social Psychology Review
1998, Vol. 2, No. 3, 196–217

HARKing: Hypothesizing After the Results are Known

Norbert L. Kerr
Department of Psychology
Michigan State University

This article considers a practice in scientific communication termed HARKing (Hypothesizing After the Results are Known). HARKing is defined as presenting a post hoc hypothesis (i.e., one based on or informed by one's results) in one's research report as if it were, in fact, an a priori hypotheses. Several forms of HARKing are identified and survey data are presented that suggests that at least some forms of HARKing are widely practiced and widely seen as inappropriate. I identify several reasons why scientists might HARK. Then I discuss several reasons why scientists ought not to HARK. It is conceded that the question of whether HARKing's costs exceed its benefits is a complex one that ought to be addressed through research, open discussion, and debate. To help stimulate such discussion (and for those such as myself who suspect that HARKing's costs do exceed its benefits), I conclude the article with some suggestions for deterring HARKing.

HARKING ET AL.

HARKing

- There is nothing wrong with constructing hypotheses after examining results.
- But: HARKing should be acknowledged explicitly as these hypotheses should be confirmed in future research.
- Because: the hypothesis is confirmed by the data upon which it was proposed. It may be statistically significant by chance, i.e., it can be a type I error (false positive).

HARKING ET AL.

HARKing

- There is nothing wrong with constructing hypotheses after examining results.
- But: HARKing should be acknowledged explicitly as these hypotheses should be confirmed in future research.
- Because: the hypothesis is confirmed by the data upon which it was proposed. It may be statistically significant by chance, i.e., it can be a type I error (false positive).

Not only HARKing...

- **Cherry-Picking** – the presentation of favorable evidence with the concealment of unfavorable evidence.
- **p-Hacking** – the relentless analysis of data with an intent to obtain a statistically significant result.
- **Fishing Expeditions, Data Dredging** – the indiscriminate testing of associations between different combinations of variables not with specific hypotheses in mind but with the hope of finding something that is statistically significant.
- ...

PRE-REGISTRATION

PRE-REGISTRATION

- **Pre-registration** (simplest form): a one page document with the basics (hypotheses, research methods and analysis plan) published prior to conducting the study.
- **Registered report**: an entire paper (without results and discussion) submitted to a journal for review and preliminary acceptance prior to conducting the study.

OTHER...

Table 1. Collection of problems and possible solutions

Section	Problems	Solutions
II.1	• Small sample size (e.g. data hard to obtain)	• Acknowledge preliminary nature • Multi-laboratory collaborations
II.1	• Novelty seeking	• Regard 'surprising' findings sceptically prior to replication
II.2*a*	• Multiple testing and selective reporting (e.g. due to too much trust in hypotheses, hindsight bias, pressure from referees)	• Avoid excessive testing (think before data exploration) • Keep track of number of tests conducted and report all tests • Bonferroni correction, false-discovery rate or emphasize preliminary nature of findings • Average effect sizes across conceptually similar tests • Referees and editors promote comprehensive and unbiased reporting
II.2*b*	• Multiple testing within models (stepwise model simplification)	• Report the initial full model • Global test of full model against null • Test a pre-determined subset of models • Average effects of individual variables across models
II.2*b*	• Overfitting of models (inflated significance)	• Keep $N > 3k$ for correct P-values, where k is number of parameters to be estimated ($N > 8k$ for reliable parameter estimates)
II.2*c*	• HARKing (hypothesizing after the results are known) and hindsight bias	• Preregister hypotheses • Keep track of number of tests conducted • Comprehensive reporting
II.2*d*	• Data collection ends with reaching $P < 0.05$	• Declare stopping rule • Adjust P-value for multiple testing
II.2*d*	• Discarding 'unsuccessful' experiments until an experiment 'works'	• Complete reporting of all experiments

LET'S PRACTICE!

IT'S TIME FOR YOUR FIRST STUDY PLAN!

KEEP CALM AND ASK QUESTIONS!

GEIST Research Group: https://geist.re/
Krzysztof Kutt: https://krzysztof.kutt.pl/