How to think about and model effect sizes

Presented by *Laura Sità*

Supervised by *D. Lakens* and *G. Altoé*

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Master thesis project

May 17th , 2024

Outline

01 Introduction

A measure of **effect size** is
a quantitative description of the strength of a phenomenon

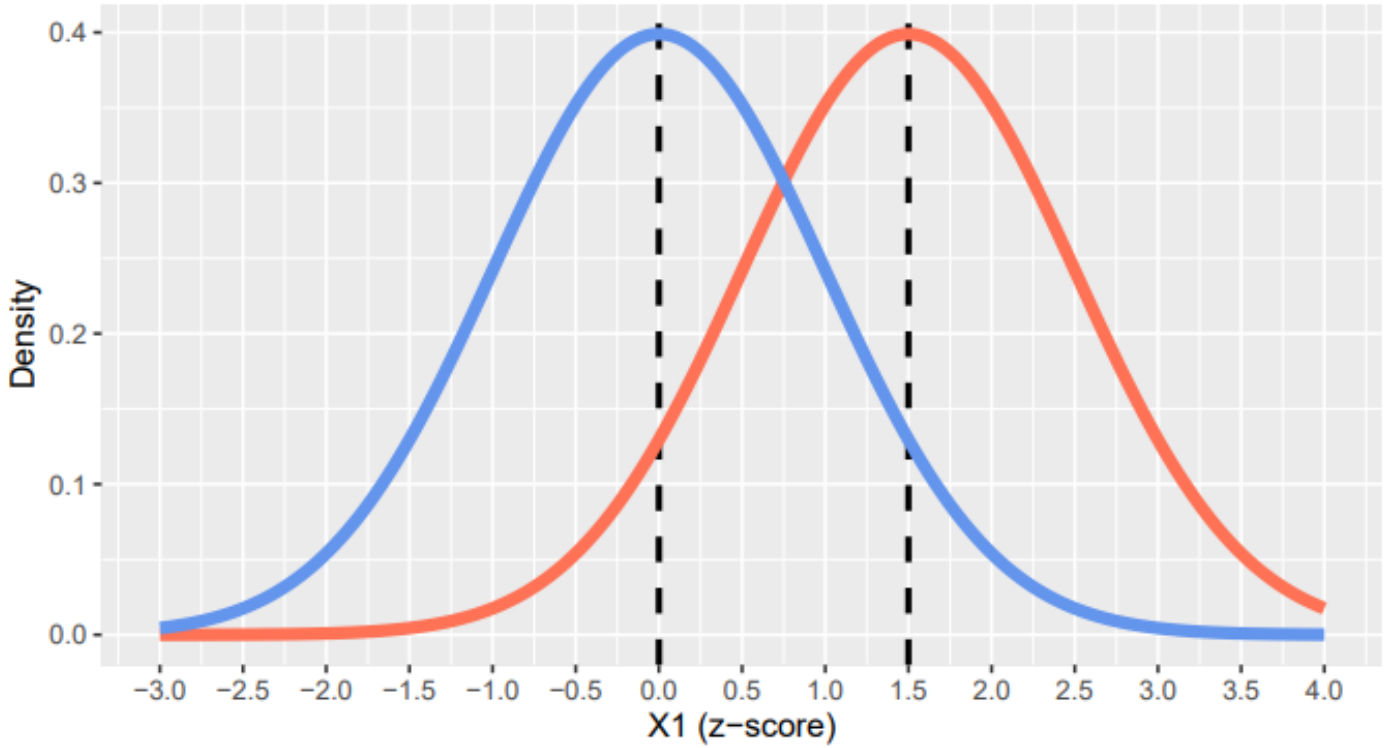
It indicates whether an intervention or experimental manipulation has an effect greater than zero or how big the effect is

For **unstandardized effect sizes**,
the effect size is expressed on the scale that the measure was collected on

To facilitate a comparison of effect sizes across situations
where different measurement scales are used

Standardized effect sizes,
which can be grouped in two families:
- the *d* family (consisting of standardized mean differences)
- the *r* family (measures of strength of association)

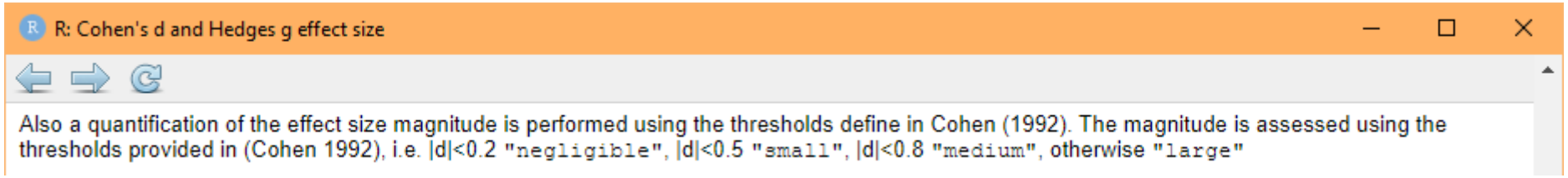
Here is an example of a visualization of two groups (“population blue” and “population pink”) representing Cohen's *d* = 1.50



Effect sizes are underappreciated and often misinterpreted

Common mistakes are to describe the effect size in ways that are uninformative
(e.g. arbitrary standards set by Cohen)
small, medium, large... compared to what/for what purpose?

Looking up the function *cohen.d* of the r-package {effsize} on the Help:



Some authors have recently suggested new solutions

HILGARD (2021)
Some effect sizes are “too large”
Suggestion. The “largest plausible effect size” is an empirical estimate of how big is too big: effects that exceed this limit are implausible and should be viewed with skepticism

FUNDER AND OZER (2019)
Effect sizes are often misinterpreted: the most common mistakes is to describe them in ways that are uninformative (e.g., using arbitrary standards) or misleading
Suggestion. Effect sizes can be usefully evaluated by comparing them with well-understood benchmarks

02 Research and objectives

This project aims to

1. suggest a better solution to think about effect sizes
2. show how to implement it in research

Achieving these objectives would lead to

- **a theoretically more meaningful way to interpret effect sizes**
- think about practical importance of the effect size

1. A better solution to think about effect sizes

Researchers should **explicitly model** the predicted effect of the independent variable on the dependent variable and consequently test it through their hypotheses

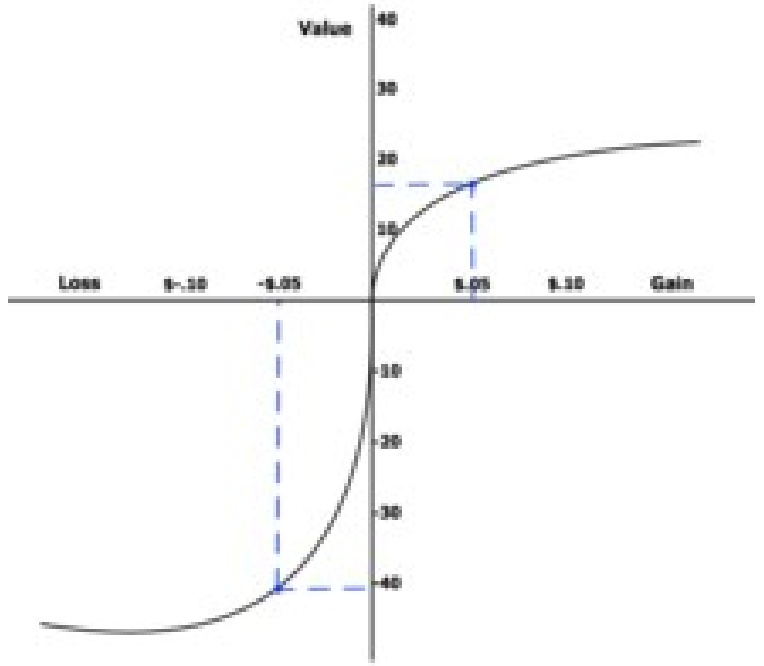
We feel that modelling would prevent researchers

- finding effect sizes too big to be true (as Hilgard pointed out)
- implicitly assuming that the relationship between variables is linear (e.g. in power analysis)

1. An example of explicitly modelling the relationship

In order to test their prospect theory, Tversky and Kahneman proposed a model showing the intensity of the manipulation as a s-shaped and asymmetrical function

They hypothesized that the value function is steeper for losses than gains indicating that losses outweigh gains



2. How to implement models in research

When it comes to design a study, we suggest psychologists to

- think about variables
- research the literature in order to find what could be the effect of their manipulation
- **create a model representing the expected effect size**
- formulate an hypothesized effect size that tests the model

Once having (simulated and then) analyzed the data

- compare the initial model (from expectations) to the one that describes the observed effect size
- interpret the comparison (e.g. if they differ, try to understand why)

03 Methodology

Data simulation using the statistical software R

1. Manipulation check

- Generate two vectors of 1000 random numbers with a normal distribution *(e.g. mood before and after a manipulation)*
- Calculate the manipulation check $\boldsymbol{x}$ from the difference between the distributions, which represents the efficacy of the manipulation *(e.g. did the manipulation affect the mood?)*
- Determine Cohen's d *(e.g. how much the manipulation affected the mood)*

2. Dependent variable

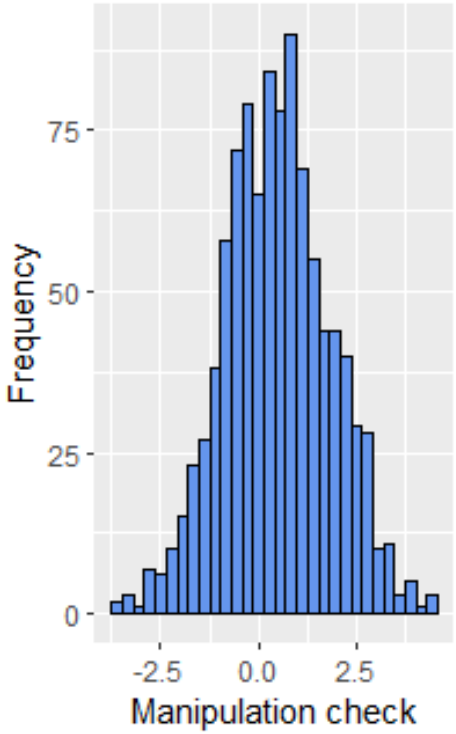
This distribution $\mathbf{y}$ is obtained from a transformation, which can be either linear or not *(e.g. how likely do you feel do make a donation to charity)*

3. Plot

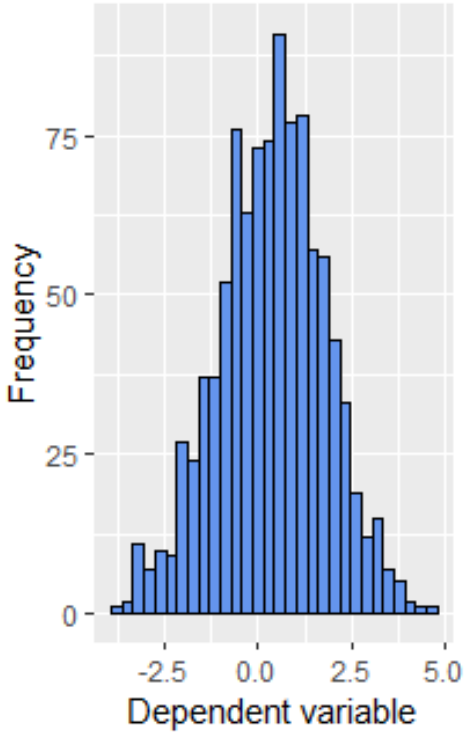
This graph shows how the effect of the manipulation is translated into an effect of the dependent variable

Example 1. Linear relationship $y = 0.8x + \varepsilon$ with $\varepsilon \sim N(0;1)$

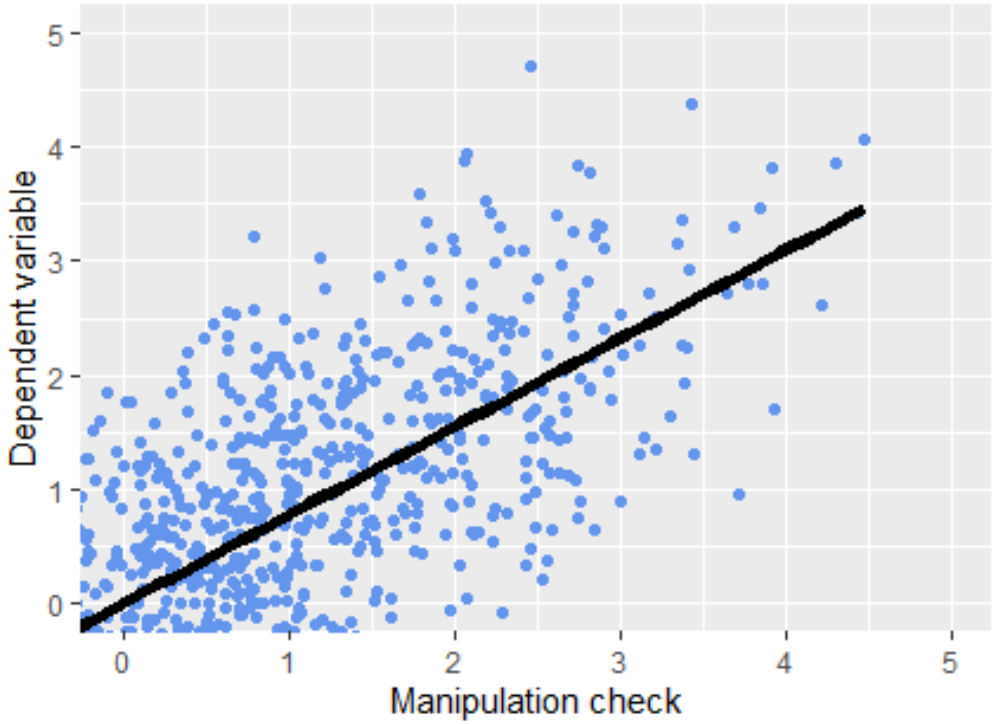
m = 0.474
sd = 1.353
d = 0.473



m = 0.359
sd = 1.434
d = 0.25

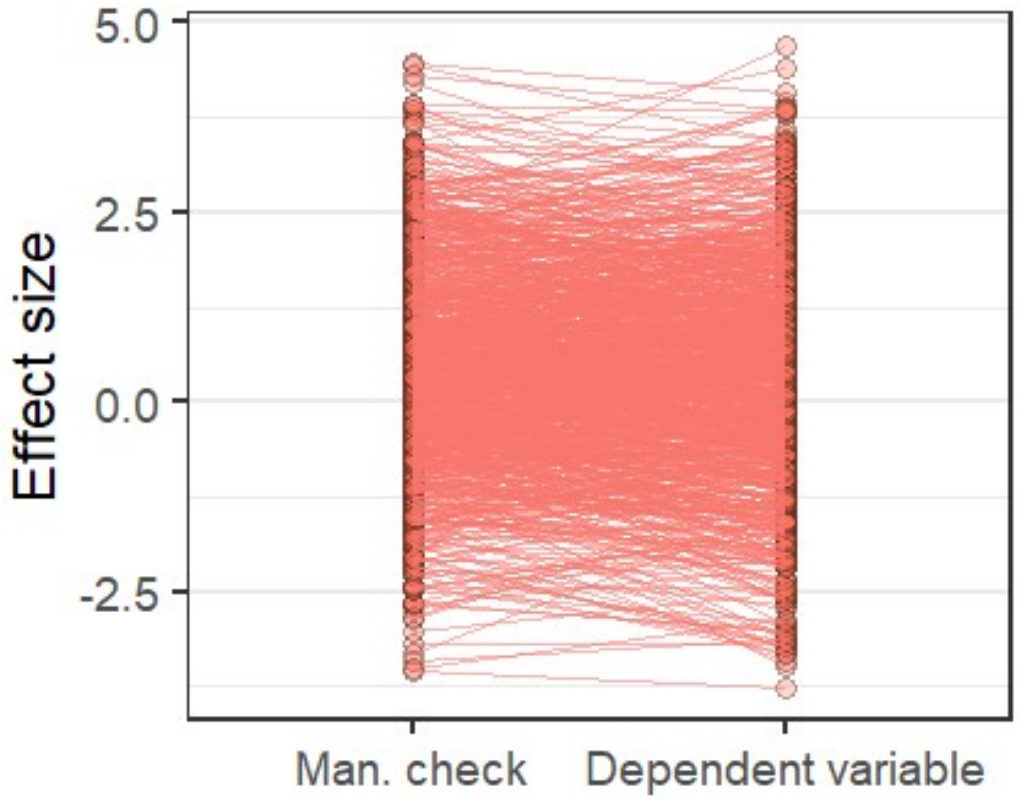


Effect of x on y



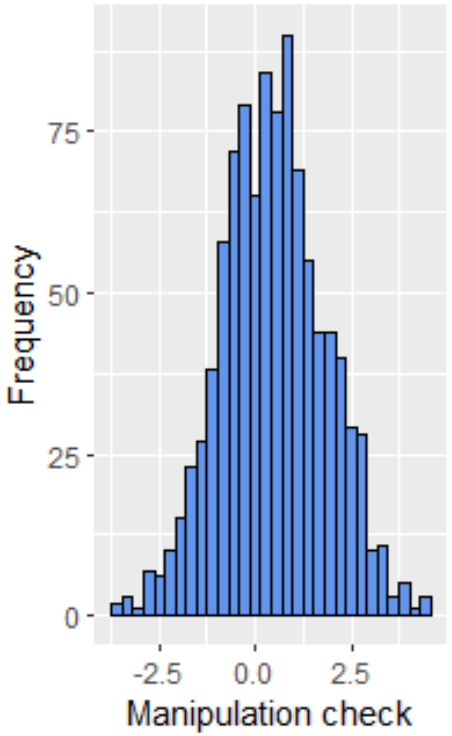
Example 1. Linear relationship $y = 0.8x + \varepsilon$ with $\varepsilon \sim N(0;1)$

How the manipulation is translated into an effect of the dependent variable

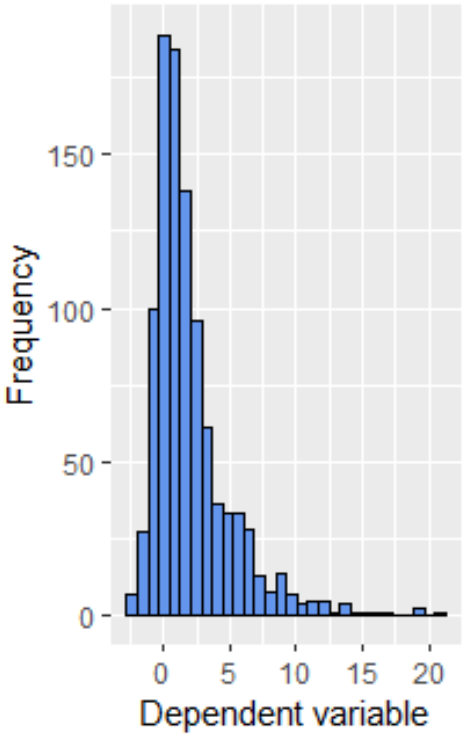


Example 2. Non linear relationship $y = x^2 + \varepsilon$ with $\varepsilon \sim N(0;1)$

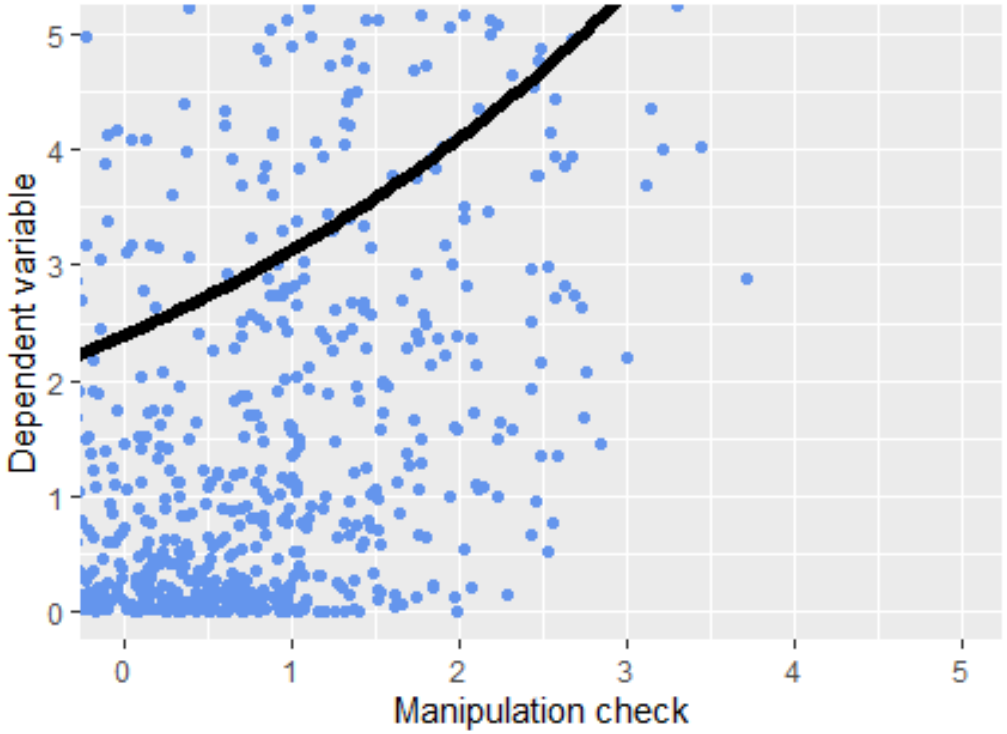
m = 0.474
sd = 1.353
d = 0.473



m = 2.032
sd = 3.016
d = 0.674

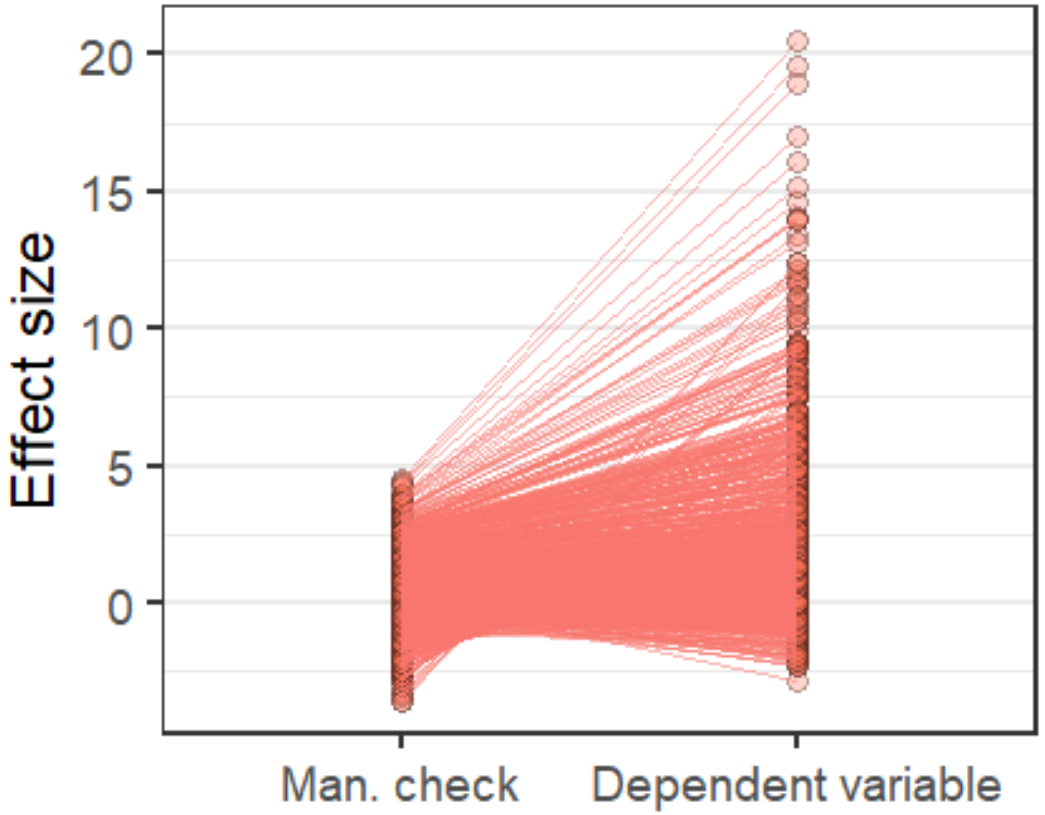


Effect of x on y



Example 2. Non linear relationship $y = x^2 + \varepsilon$ **with** $\varepsilon \sim N(0;1)$

How the manipulation is translated into an effect of the dependent variable



Empirical component of the project: analysis of a real data set

1. Find a open-access study where
- data are accessible
- a manipulation is done and a manipulation check is measured
- the dependent variable is related to the manipulation through some psychological mechanism

2. Create a plot that
- highlights the relationship between the two distributions of data (subjects' responses for the manipulation check and the dependent variable)
- represents the links between each couple of responses within every subject

3. Expect to find
- a correspondence within each couple of data
- that would be graphically represented by a series of straight lines

We are looking for studies
with an **available data set** where the researchers conducted

- a manipulation on the subjects and collect data to check if it happened as planned (**manipulation check**)

- a **measurement of the dependent variable**, related to the manipulation through some psychological mechanism

04 Conclusions

Key points

→ This work is a step forward to a meaningful interpretation because it suggests

- an alternative way to think about expected effect sizes
- to explicitly model the relationship when designing their study makes researchers think thoroughly about the effect they are studying

→ This procedure can be applied within both the frequentist and bayesian approach

Future research

- → Work on setting lower and upper bounds (like Hilgard suggested with estimate of the largest plausible effect size)
- → Spread the procedure of explicit modelling with researchers in the various fields of psychology (e.g. predict the effect of a clinical intervention)
- → Anything else ?

The code is available on RPubs
https://rpubs.com/laurasita/1185335

Thank you for your attention

Any questions or feedback?

Email address:
laura.sita@studenti.unipd.it