

UNIVERSITY OF TWENTE.



RESEARCH METHODS

DASHA KOLESNYK



WHAT ARE WE DOING HERE?

WE ARE LEARNING A TRANSFERRABLE SKILL



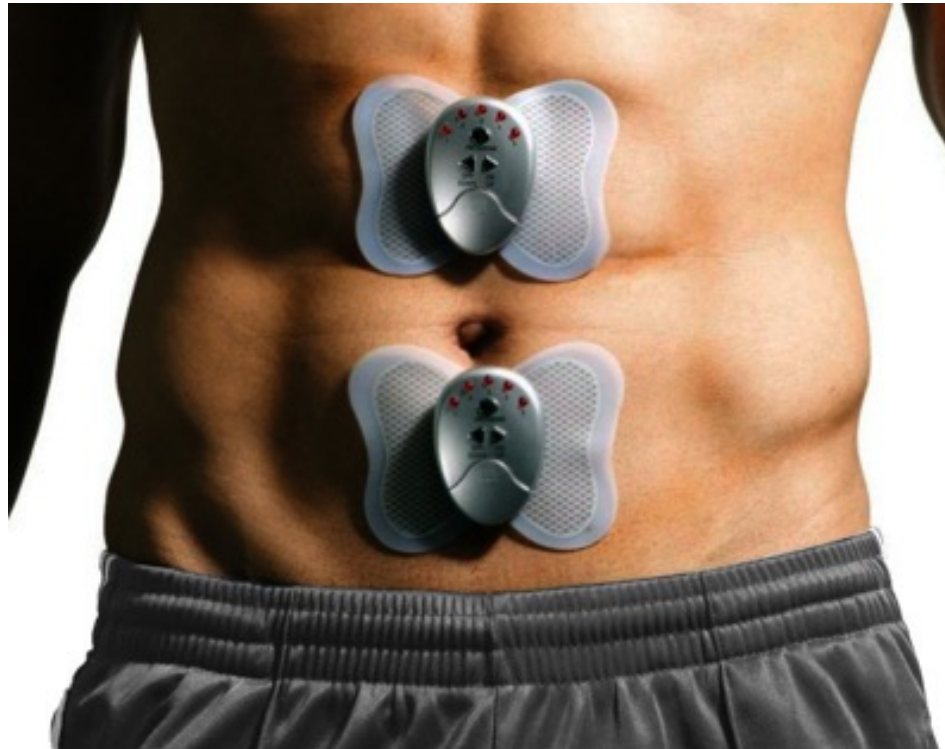
How to know whether a causal statement is true?

Should you believe what the product description says?

How can you design your story in a way that people not only notice it, but also BELIEVE it?

EXAMPLE

WHAT WOULD IT TAKE FOR YOU TO BE SURE THAT IT WORKS?



WHY KNOW ANYTHING ABOUT SCIENTIFIC METHODS?

TAKE IT AS A GENERAL GUIDANCE INTO CRITICAL THINKING

NOT the answer to all of your problems.

Currently best way we know to be as sure as possible that what we think happens is what actually happens.

In HMI we use it when we try to prove that it works as we say.

RESEARCH METHODS

(ONE OF CLASSIFICATIONS)

- Descriptive (X and Y happen)
- Relational (X and Y change together)
- Experimental (change in X **causes** the change in Y)

Example from the book:

relationship between using computers and typing.

It's all about details

CAUSALITY

#typicalteacher



Correlation
DOES NOT
imply
causality!

FOCUS ON EXPERIMENTAL RESEARCH



When we want to know whether
change in the independent variable (X)
CAUSES
change in the dependent variable (Y)

INDEPENDENT VS DEPENDENT VARIABLES

Independent variable (IV) **X**

The factor, cause of the change in the dependent variable
The treatments or conditions the researcher can control

Dependent variable (DV) **Y**

Outcome or effect the researcher is interested in
The outcome the researcher needs to **measure**

Let's have a small
terminology practice

FORMULATE RESEARCH QUESTION

WHAT IS THE QUESTION YOU WANT AND CAN ANSWER

Be SPECIFIC!

Why/How using computers affects typing? – too broad, no specific answer

Do people who use computers more get faster at typing? – answerable

How much the speed of typing improves with increase in time of computer use?

FORMULATE YOUR HYPOTHESIS (OR HYPOTHESES)

WHAT IS THE CHANGE YOU EXPECT TO SEE IN Y AS YOU CHANGE X?

Null: there is no difference between two or more conditions

Alternative: is always mutually exclusive with the null hypothesis, there is a difference

Be SPECIFIC! 😊

Null: Typing speed will NOT change with use of computers

Alternative: Typing speed will change with use of computers

Null: Typing speed will not change as people use computers more time

Alternative: Typing speed will increase as people use computers more time

Null: After two weeks of using computers 3 hours more per day than usual typing speed will remain the same

Alternative: After two weeks of using computers 3 hours more per day than usual typing speed will increase by at least 20 words per minute

Let's have a
small
formulation
practice

EXPERIMENTAL RESEARCH

YOU ONLY CHANGE ONE THING AT A TIME, EVERYTHING ELSE REMAINS EQUAL



EXAMPLE RESEARCH AT HMI (ZAGA ET AL., 2015)

OPERATIONALIZED AND MEASURABLE, AND NEUTRAL

Which robot, a peer-like or a tutor-like, will yield better task efficiency (task completion and duration) for children?



EXAMPLE HYPOTHESIS

RQ: Which robot, a peer-like or a tutor-like, will yield better task efficiency (task completion and duration)* for children?

H0: There is no difference in children's task efficiency (task duration and completion) between a peer-like and a tutor-like robot

H1: Children's task efficiency (task duration and completion) is higher with a peer-like robot

**if you cannot measure it, we cannot talk about it...*

Let's have a
measurement
practice

STUDY DESIGN

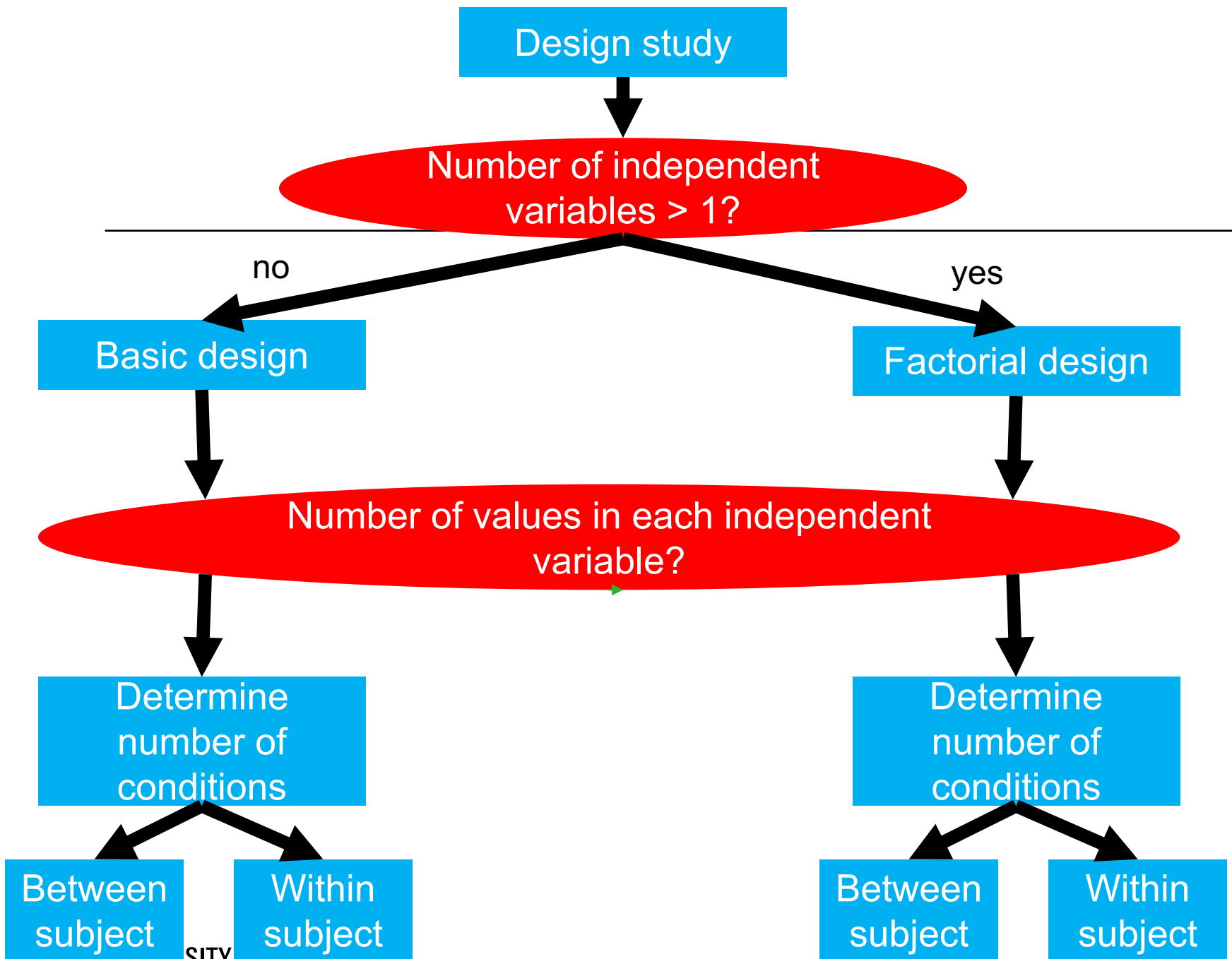
SO WE HAVE A RESEARCH QUESTION, NOW WE DECIDE HOW TO TEST IT

How do I answer the research question in an actual study?

What do I want to compare against what?

How do I design a study in such a way

- 1) that I can trust the results, and
- 2) that the results are answering the research question?



EXAMPLE INDEPENDENT AND DEPENDENT VARIABLES

RQ: Which robot, a peer-like or a tutor-like, will yield better task efficiency (task completion and duration) for children?

IV: What factor to control (manipulate) and how many values?

Character of the robot, 2 values: peer and tutor

DV: What outcome to measure?

Task completion (in %)

Task duration (in seconds)

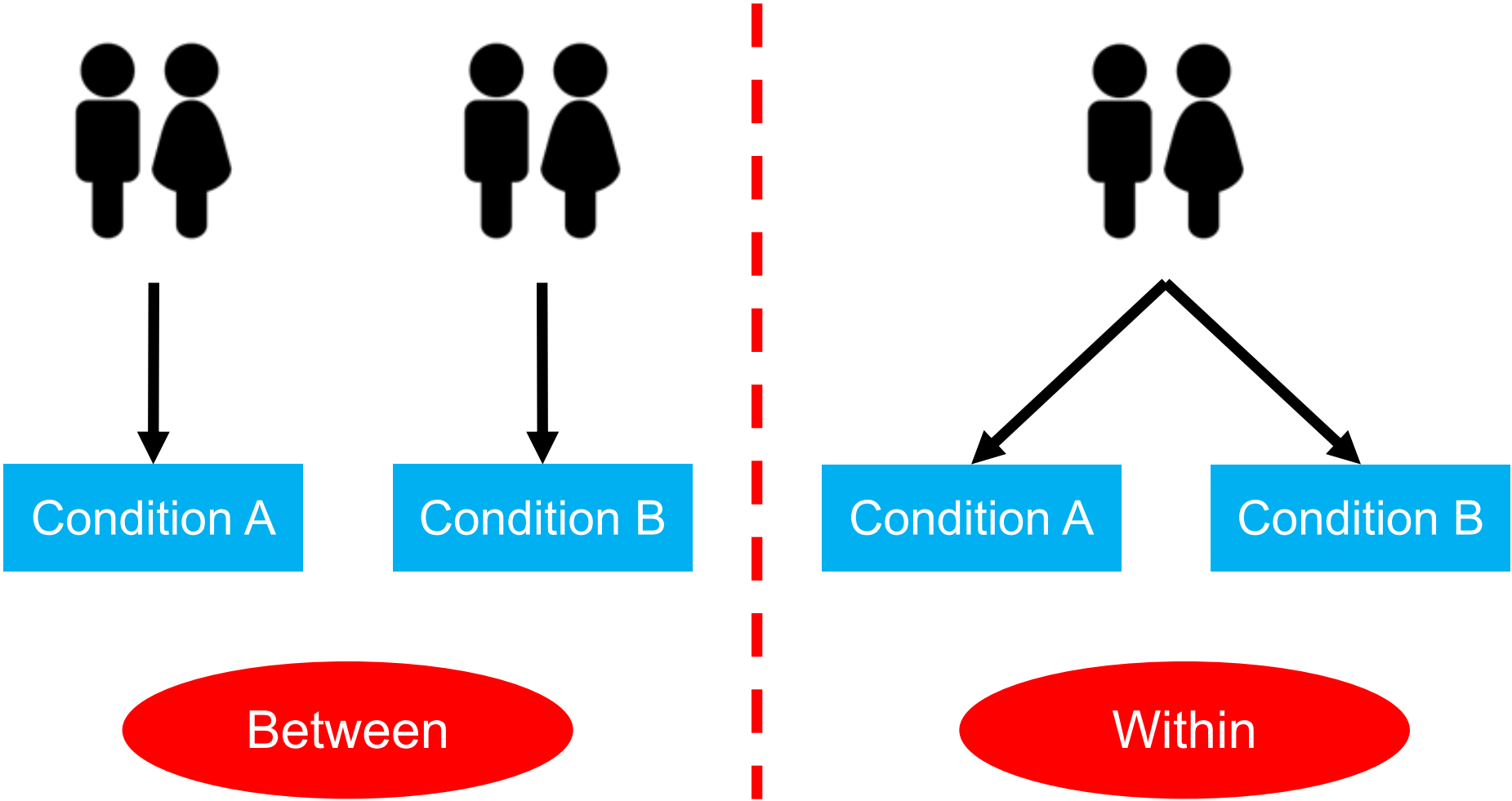
STUDY DESIGN: NUMBER OF CONDITIONS

2 conditions: tutor-like vs peer-like



Need to decide whether 1) all children interact with both tutor-like and peer-like robots, or 2) group A only interacts with tutor-like and group B only interacts with peer-like

BETWEEN VS WITHIN SUBJECT DESIGN



WITHIN

YES, BUT....

Advantages

Smaller sample size

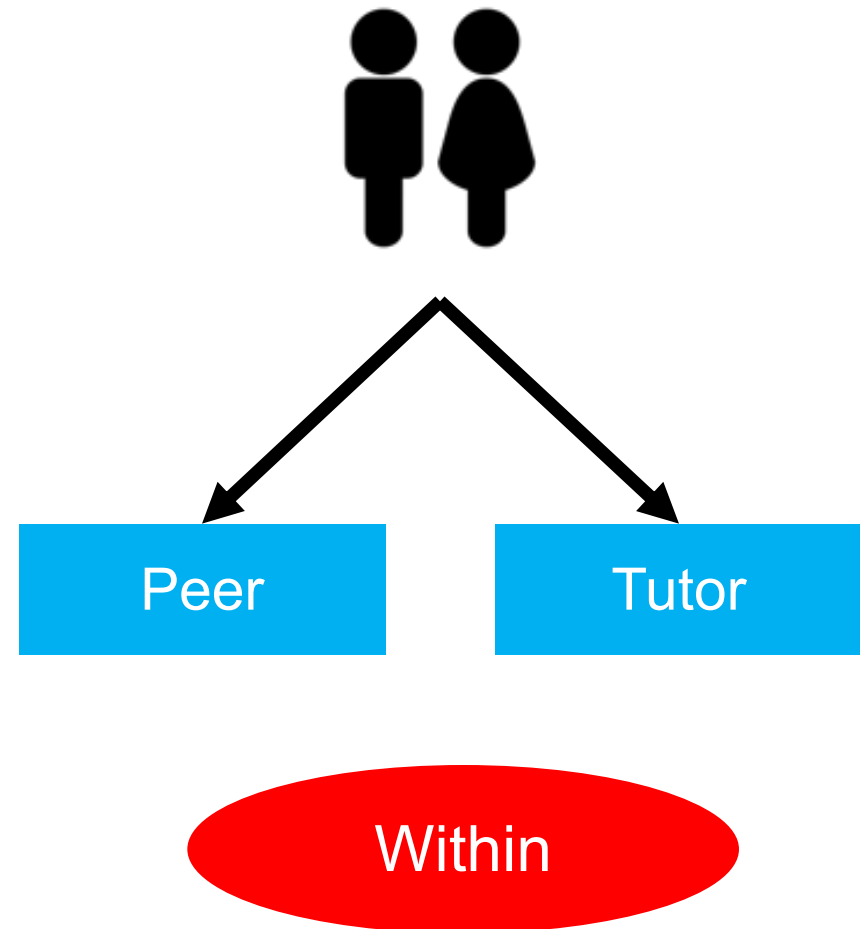
Less impact individual differences

Disadvantages

Hard to control learning effect

Impact of fatigue

Watch out!



BETWEEN

YES, BUT....



Peer



Tutor

Advantages

Avoids learning effect

Better control of e.g., fatigueness

Disadvantages

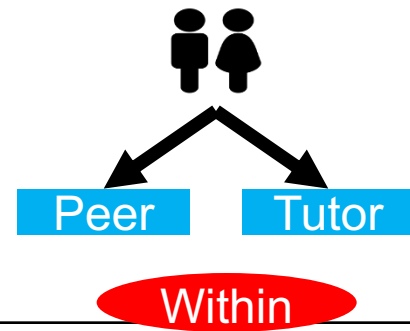
Need larger sample size

Larger impact individual differences

Between

PITFALLS - WITHIN

YES, BUT....



Learning effects – performance gets better with experience, learning curve not the same for everyone

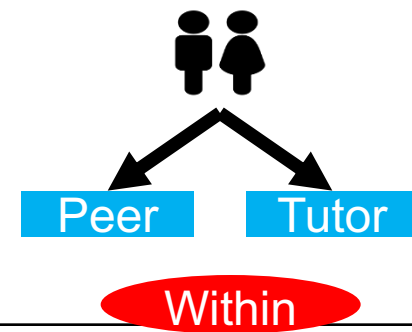
✓Provide sufficient time for training

Novelty effect - some people have never seen a robot before

✓Provide sufficient time for getting used to the robot

PITFALLS - WITHIN

YES, BUT....



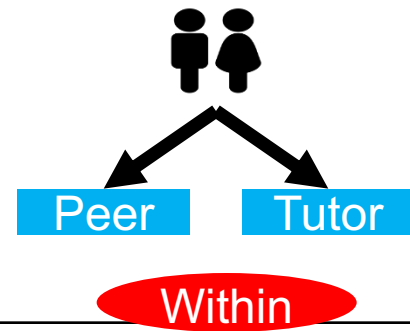
Order effect – order of conditions may matter and influence results

✓ **Counterbalancing:** each possible sequence is used $n!$, e.g. $3!=6$ possible sequences

	1	2	3
1	A	B	C
2	A	C	B
3	B	A	C
4	B	C	A
5	C	A	B
6	C	B	A

PITFALLS - WITHIN

YES, BUT....

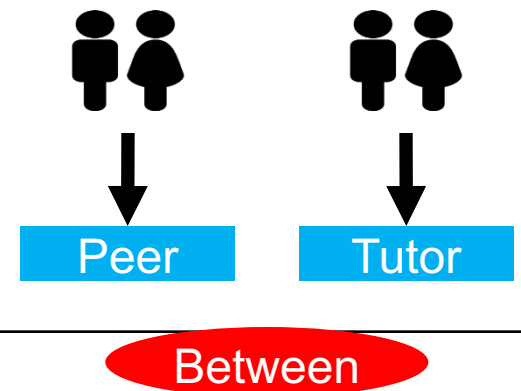


Fatigue

- ✓ Do not make the task too long
- ✓ Add some breaks if necessary

PITFALLS - BETWEEN

YES, BUT....



Larger impact individual differences

- ✓ Choose simpler tasks, limited cognitive process
- ✓ Participants must be **randomly** assigned to different conditions, the participant groups should be as similar as possible

Let's
Practice designing a
study

WHAT IF YOU HAVE 2 INDEPENDENT VARIABLES?

How long would children pay attention to the robot? (time in seconds measured from start to first non-compliance)

	Tutor-like	Peer-like
Says a joke	250 sec	360 sec
Doesn't say a joke	247 sec	300 sec

Invented example!

STATS



STATISTICS AS AN ART

**SIGNIFICANCE TESTING – WHY
SCALES
T-TEST
ANOVA
REGRESSION**

WHY STATISTICS?

WOULD YOUR ANSWER CHANGE IF I TOLD YOU THAT THIS IS BASED ON 2 CHILDREN ONLY?

How long would children pay attention to the robot? (time in seconds measured from start to first non-compliance)

	Tutor-like	Peer-like
Says a joke	250 sec	360 sec
Doesn't say a joke	247 sec	300 sec

Is there a difference?

WHY STATISTICS?

How long would children pay attention to the robot? (time in seconds measured from start to first non-compliance)

	Tutor-like	Peer-like
Says a joke	286 sec	360 sec
Doesn't say a joke	247 sec	300 sec

Is there a difference?

**SO HOW DO WE KNOW WHETHER THERE IS A
DIFFERENCE?**

**STATISTICAL ANALYSIS:
SIGNIFICANCE TESTING**

So many tests, so little
time...

How to choose?

SOFTWARE

SPSS, SAS, Stata, Excel, Matlab, **R...**

INTERNET

WHAT YOU NEED TO KNOW TO SEARCH WHICH STATISTICAL TEST YOU NEED?

“[type of scale] dependent variable [type of scale] independent variable
which test”

Types of scales

Nominal

Ordinal

Continuous

Interval

Ratio

WHAT TYPE OF VARIABLES ARE THESE?

Where are you from?

- Netherlands Outside NL, within Europe Outside Europe

How old are you?

- Less than 20 21-22 23 - 24 25 and above

I love statistics

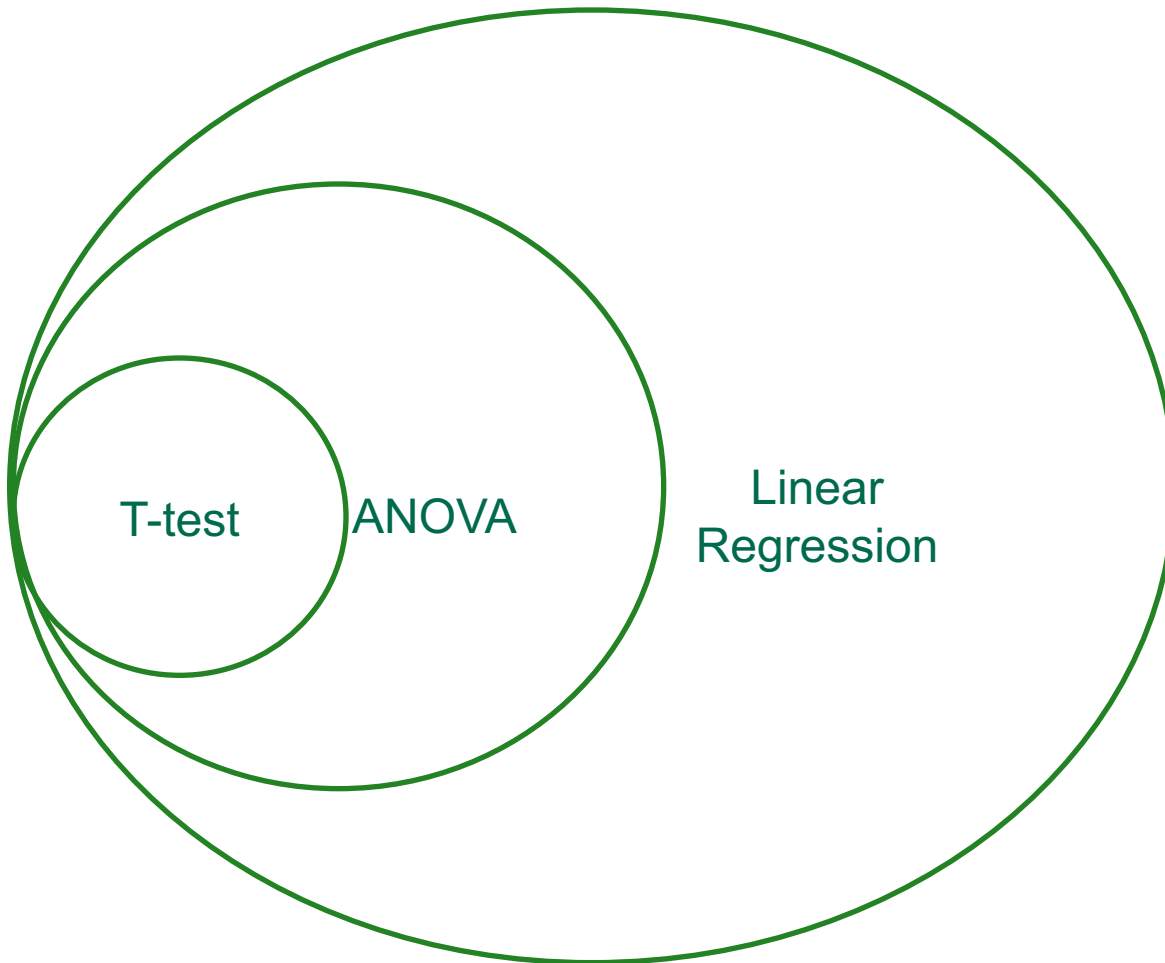
- | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Strongly | Somewhat | Neither agree | Somewhat | Strongly |
| Disagree | disagree | nor disagree | agree | agree |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

The grade I expect to receive in the course: _____ (out of 10)

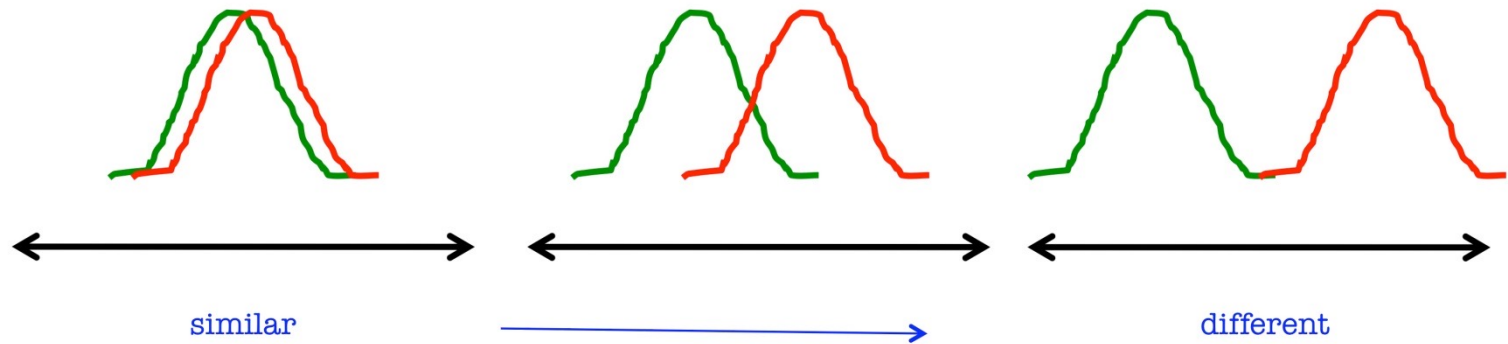
MULTIVARIATE STATISTICAL METHODS

		Independent Variable		
		Nominal	Ordinal	Interval/ratio
Dependent Variable	Nominal	Chi-squared test	Choice model	Choice model
	Ordinal	Mann-Whitney	Spearman-rank correlation	Ordered choice model
	Interval/ratio	t-test, ANOVA Linear regression	Linear regression	Linear regression

INTUITION!



T-TEST: ARE TWO MEANS STATISTICALLY DIFFERENT?



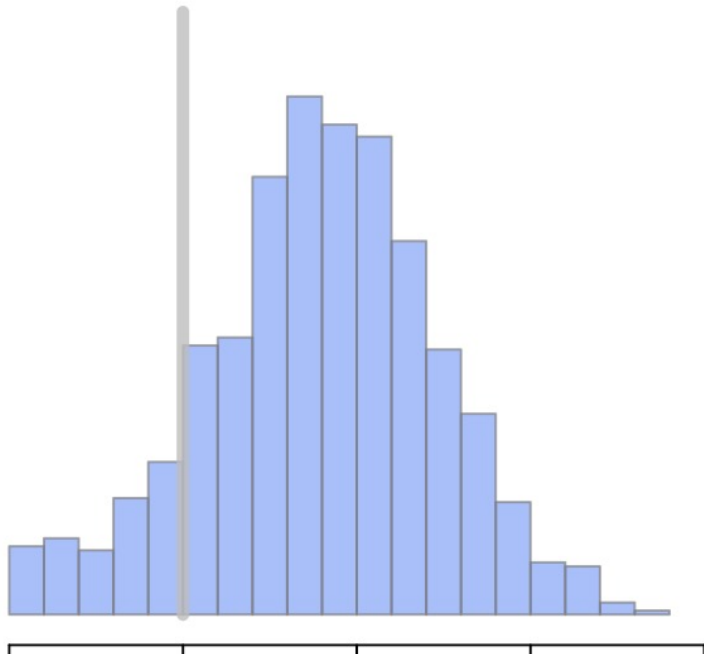
ONE SAMPLE

T-TEST

VS

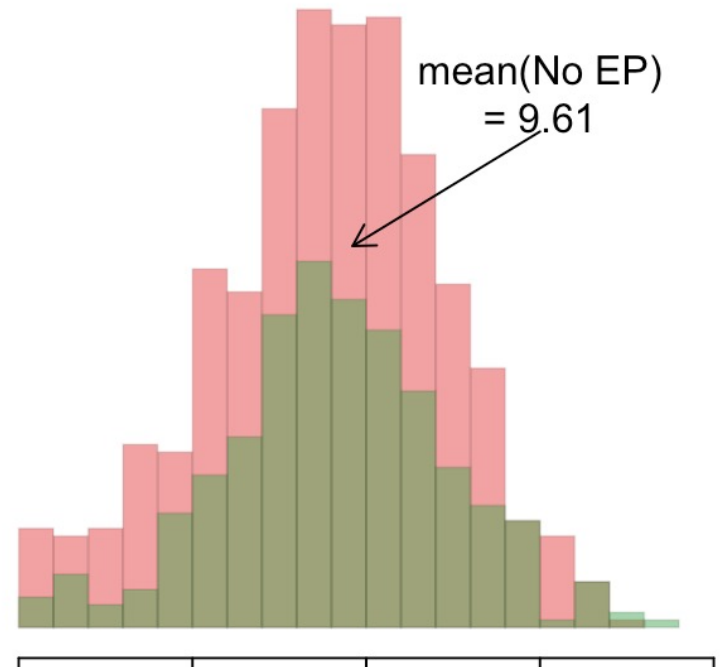
TWO SAMPLE

Null Hypothesis
Mean = 5



mean(EP)
= 9.34

mean(No EP)
= 9.61



T-TEST: ARE TWO MEANS STATISTICALLY DIFFERENT?

One-tailed vs. two-tailed

Example: Greater/less than (one-tailed) or significantly different (two-tailed)

Depends on clarity of
your hypothesis

t-test

- **One sample t-test**

Comparing data across several observations with a single number

Example: Is the mean age of this class more than 21?

- **Independent sample t-test**

Comparing data across several observations for two independent groups

Example: Will this class perform better than that of last year?

- **Paired sample t-test**

Comparing data for the same observations in two settings

Example: Will this class perform better in Branding course when compared to MRA?

SPSS

TIP: "SPSS ANNOTATED OUTPUT" ONLINE

		t-test for Equality of Means						
		t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Cm loss in waist	Equal variances assumed	3.773	124	.001	4.231	.299	4.060	4.422

Probability of such mean distribution given H_0 .

$P < 0.05$ (we just agree that this is good enough)

ANOVA: ANALYSIS OF VARIANCE

FOR YOUR EXPERIMENTAL DESIGN

Depends on your
experimental design

2x1 One-way ANOVA

Control group	Experimental Group
No butterfly	Butterfly 30 min/day



Same as

- Independent samples t-test if the groups are different (between-subject design)
- Paired samples t-test if the groups are same (within-subject design)

3x1 One-way ANOVA

Control group	Experimental Group1	Experimental Group 2
No butterfly	Butterfly 30 min	Butterfly 2 hours



2x2 Two-way ANOVA

	Tutor-like	Peer-like
Says a joke	286 sec	360 sec
Doesn't say a joke	247 sec	300 sec

- Main effects: effect of an independent variable on a dependent variable averaging across the levels of any other independent variables.
- Interaction effect: Two independent variables interact if the effect of one of the variables differs depending on the level of the other variable

Linear Regression Model

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \varepsilon_i = X_i' \beta + \varepsilon_i$$



Linear regression model is generalization of traditional correlations to a multivariate setting

Example Research Question

Determining the drivers of the efficiency of teaching by the robot



Example – Model

Number of tasks solved_h = β_0 +

β_{type} robot_type_h +

β_{joke} joke_h +

β_{color} color_robot_h +

β_{gender} gender_child_h +

β_{age} age_child_h + ϵ_h



EXAMPLE OUTPUT

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	12.553	4.461		3.724	.000
	Robot type: peer-like (dummy)	4.539	2.850	.286	3.891	.023
	Makes joke (dummy)	1.110	1.935	.162	1.553	.123
	Robot red (dummy)	2.545	2.220	.089	.936	.351
	Child gender (dummy)	1.396	1.417	.038	.366	.715
	Child age (years)	1.234	0.56	.245	2.345	.012

a. Dependent Variable: number of tasks completed

B tells you how much DV changes with one unit change in IV when all other variables are at 0.

STATISTICS AS AN ART

**SIGNIFICANCE TESTING – WHY
SCALES
T-TEST
ANOVA
REGRESSION**

MUCH MORE TO DISCOVER

PLAY

SUMMARY

- Do what you love in a way that it makes you love what you do. Example?
- When do we need experimental research?
- Does correlation imply causality? 😊
- What is the main point of an experiment?
- What's a between-subject design?
- What's a within-subject design?
- Give an example of potential interaction effect.