

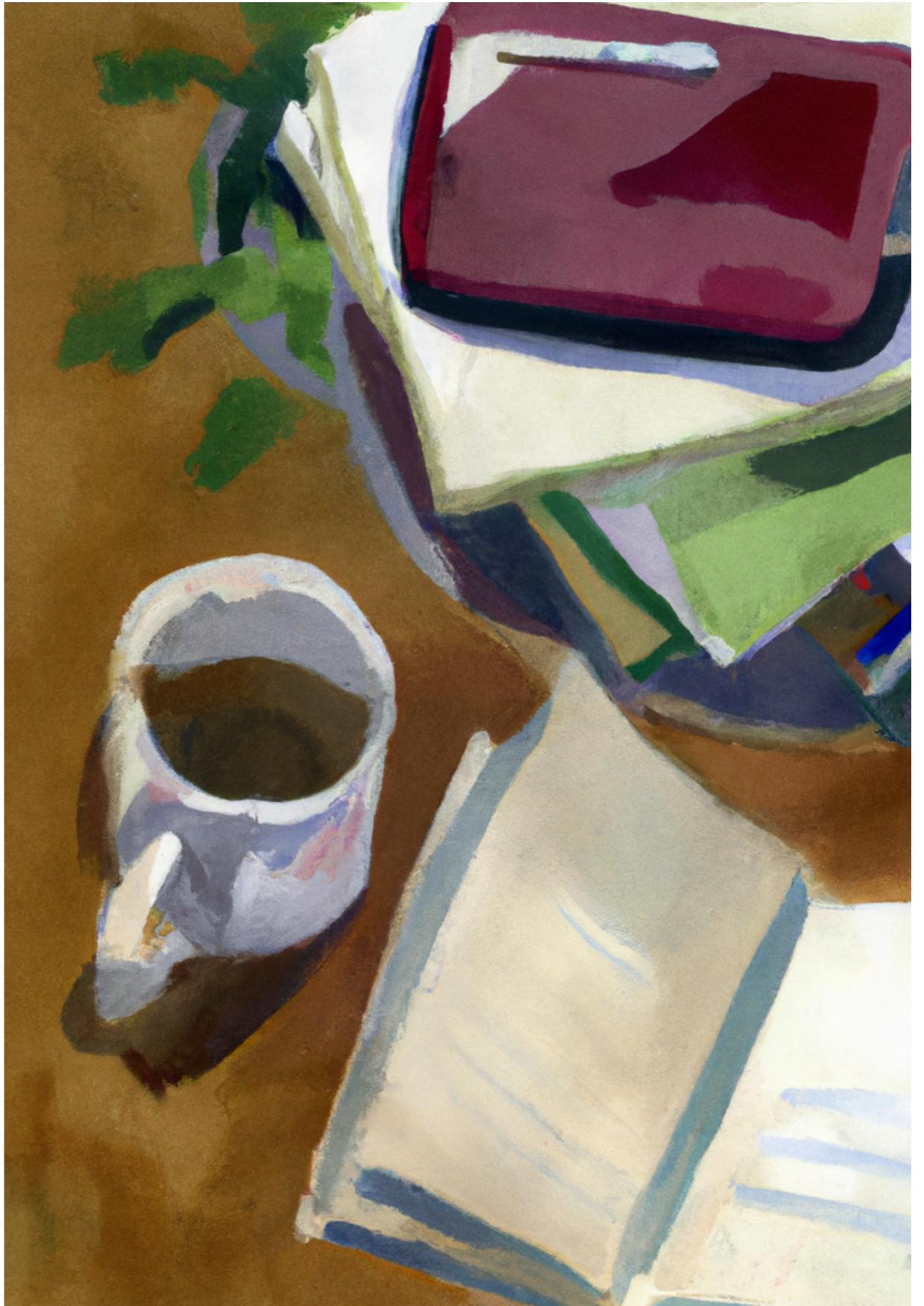


Week 2: Design Analytics

Experimental Design/2

Agenda

- **Recap**
 - Research Questions (RQs)
 - Hypothesis (Null & Alternative)
 - Independent & Dependent Variables
 - Baseline (or Ground Truth)
- **Coming Next!**
 - Control Group/Condition
 - Intervention or Treatment
 - Experiment Design
 - Errors and Biases

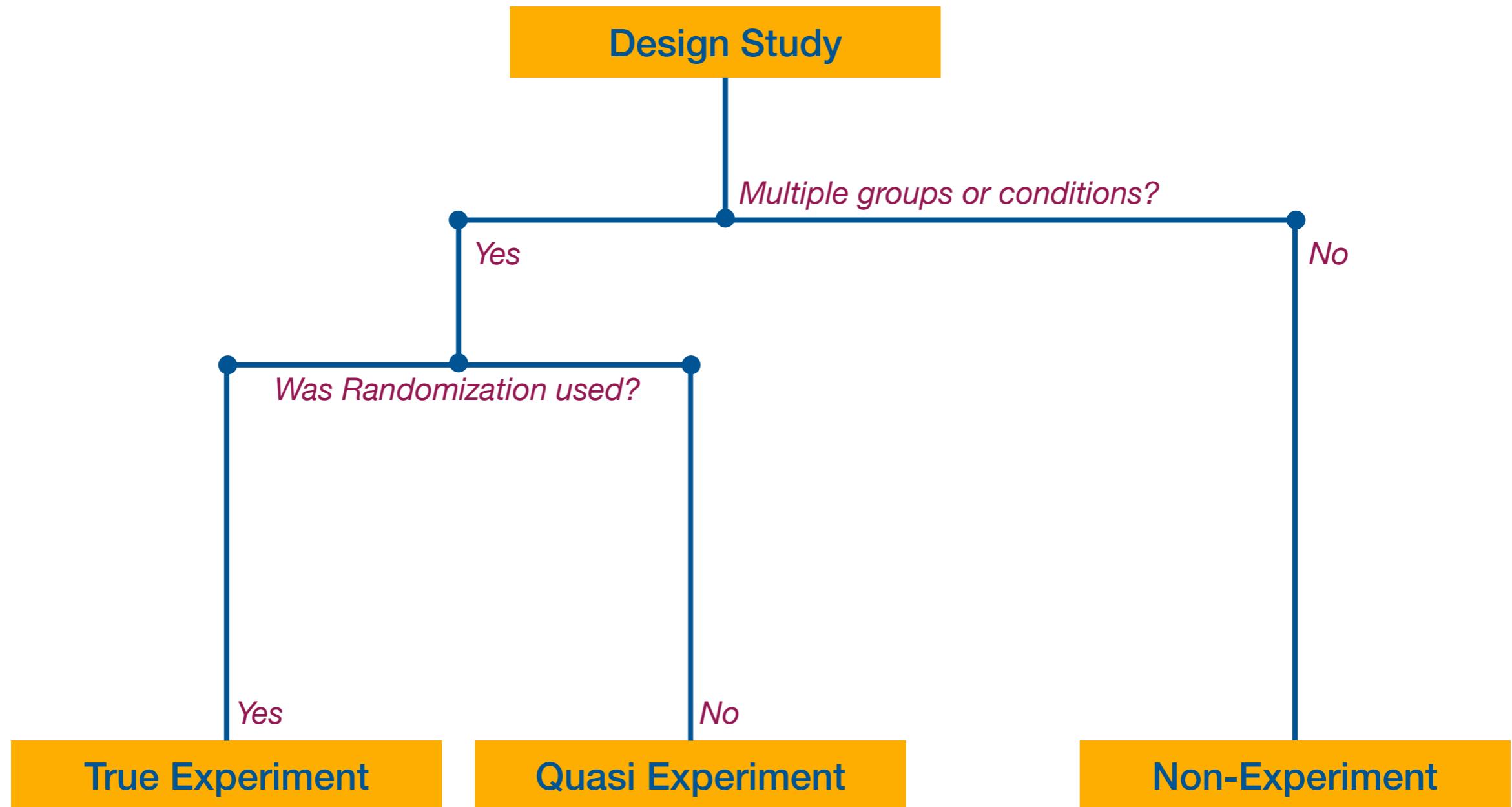


Sound Experiment: Ingredients

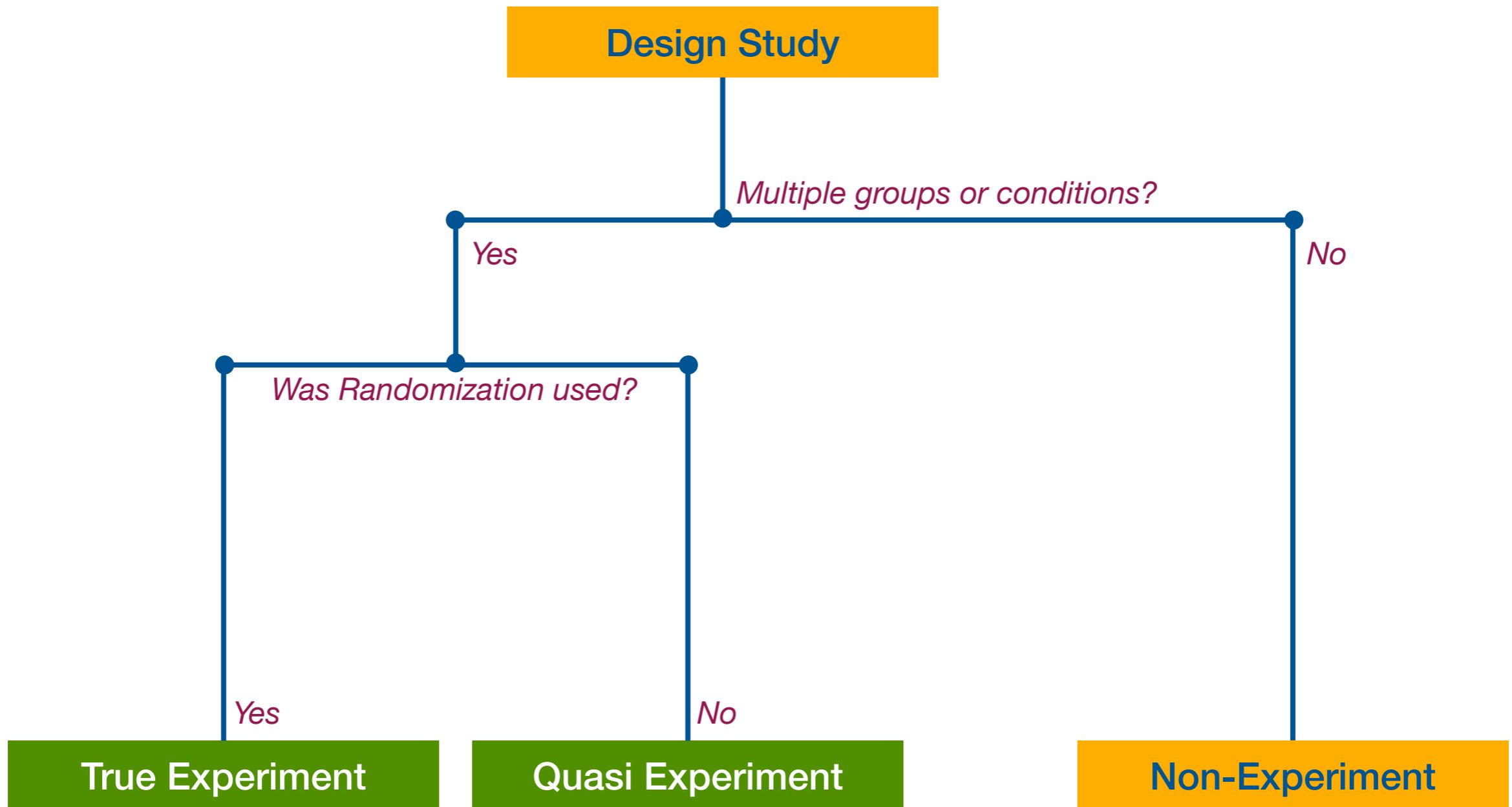
- 1x testable **hypothesis**
- 2x **experimental conditions** (*control/baseline vs. treatment/intervention*).
- Dependent Variables (DVs) can be measured **quantitatively**.
- There are **no biases** in the experiment design.
- Results can be analyzed by **statistical** means.
- Validity can be assured, i.e., from **replicability** (*across samples, times, locations, experimenters, etc.*) to **reproducibility**.
 - **Internal** validity (*reliability*)
 - **External** validity (*generalizability*)
 - **Ecological** validity (*realism*)



True, Quasi or Non-Experiments



True vs. Quasi Experiments/1



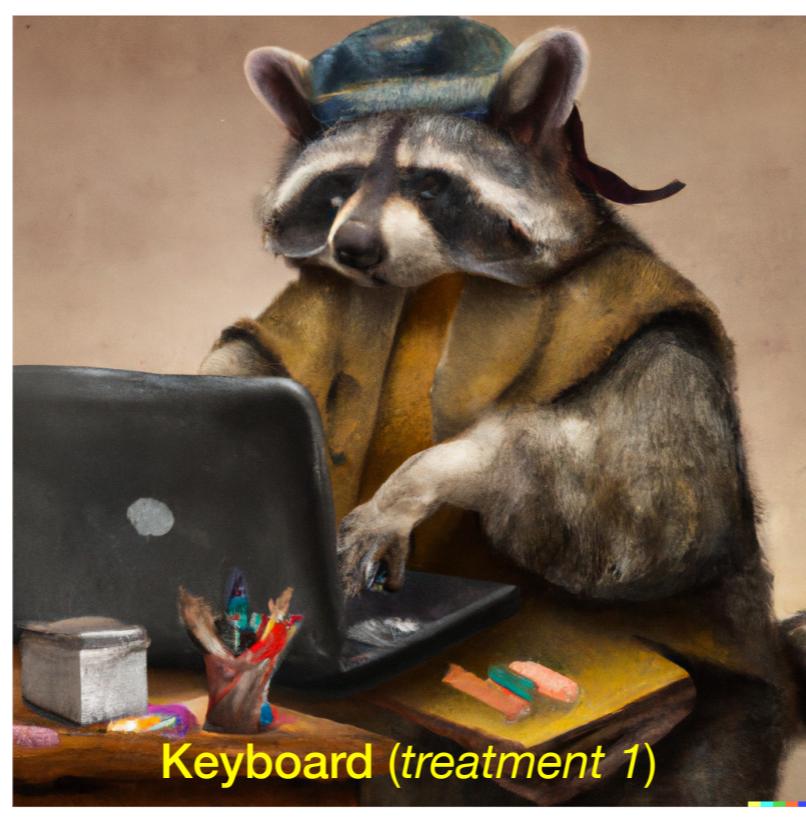
True vs. Quasi Experiments

	True Experiments	Quasi Experiments
How are participants assigned to control & experimental groups?	Randomly. Both researcher and participants are blind to assignment.	Non-Random. Participants may not be blind to assignment.
Is there control over experimental group?	Yes. Researchers design the experimental (treatment) group.	The researcher often does not have control over the treatment group, but instead studies pre-existing user groups that have received other similar treatments
Are control groups used?	Yes. It is a must to have a control (baseline) and an treatment group (at least one).	Control groups are not required, but they are still quite common.
What is the emphasis?	Internal Validity: Used to test well-defined hypothesis and assess cause-effect relationships (usually in lab settings)	External Validity: Used to test well-defined and exploratory hypothesis and investigate descriptive aspects.

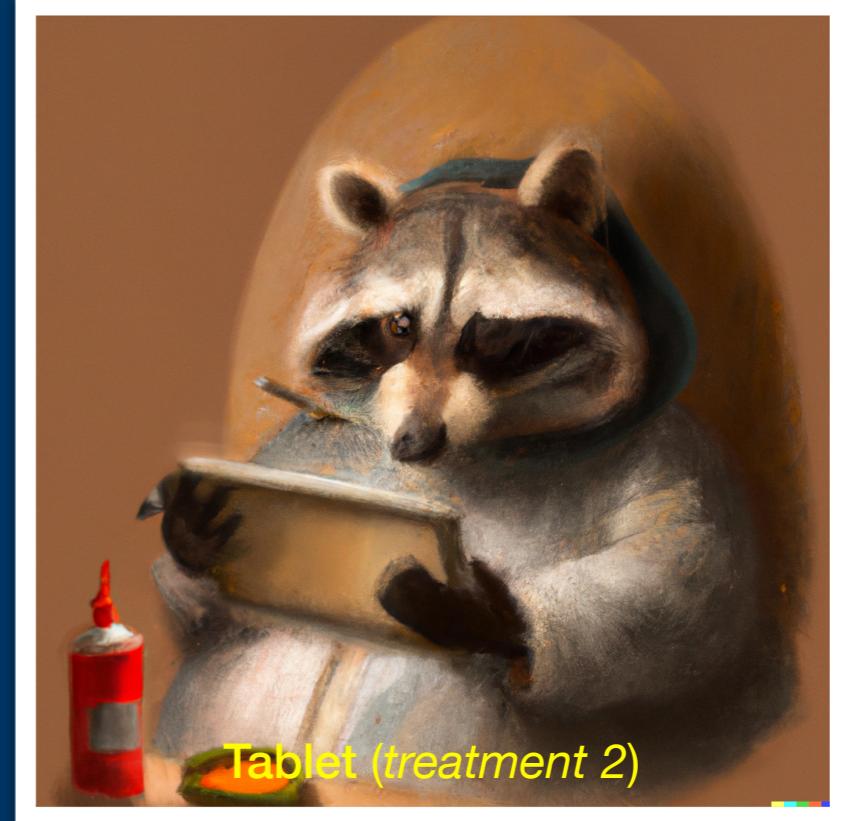
*Experimental group/condition is also referred to as treatment.



Pen (*control group*)



Keyboard (*treatment 1*)



Tablet (*treatment 2*)

Example: Writing Text

Research Questions (RQs)

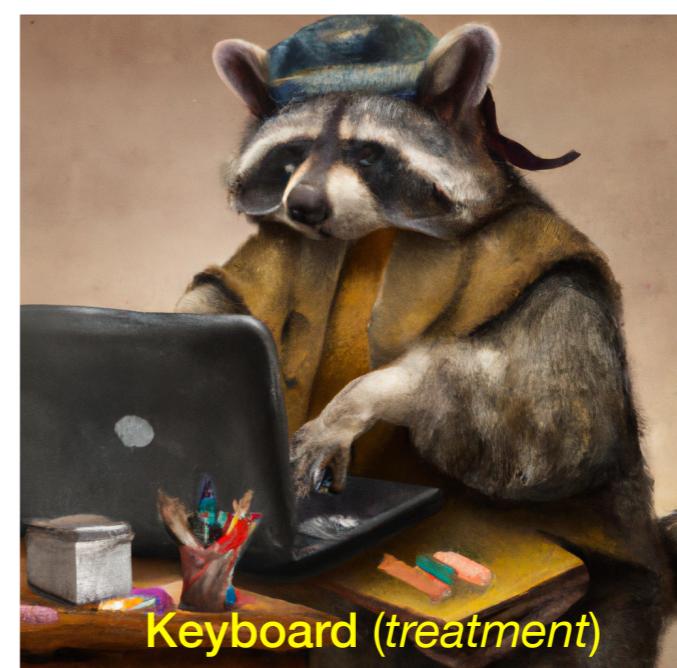
RQ1: Is a keyboard an “**effective**” way of producing **textual** content?

RQ2: Is a keyboard an “**ergonomic**” way of producing **textual** content?

RQ2: Is a keyboard a “**versatile**” way of producing **different types** of content?



Pen (*control group*)



Keyboard (*treatment*)

Hypotheses (H_0 & H_1)

RQ1: Is a keyboard an “**effective**” way of producing **textual** content?

H_0 : There is no difference between Pen and Keyboard in terms of speed of writing.

H_1 : The speed of writing text is higher for Keyboard as compared to Pen.

RQ2: Is a keyboard an “**ergonomic**” way of producing **textual** content?

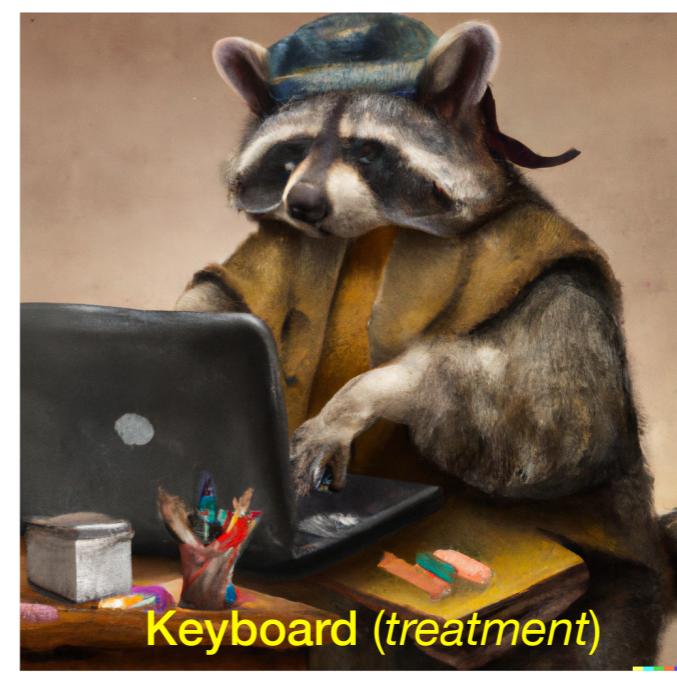
H_0 : There is no difference between Pen and Keyboard in terms of fatigue while writing.

H_1 : The fatigue while writing text is higher for Keyboard as compared to Pen.

RQ2: Is a keyboard a “**versatile**” way of producing **different types** of content?

H_0 : There is no difference between Pen and Keyboard in terms of ease of producing drawings.

H_1 : The ease of producing drawings is higher for Pen as compared to Keyboard.



Variables and Conditions

Independent Variables (IVs) or Conditions

Input Device (Pen vs. Keyboard) — *2 levels*

Control Condition: Pen (*baseline*)

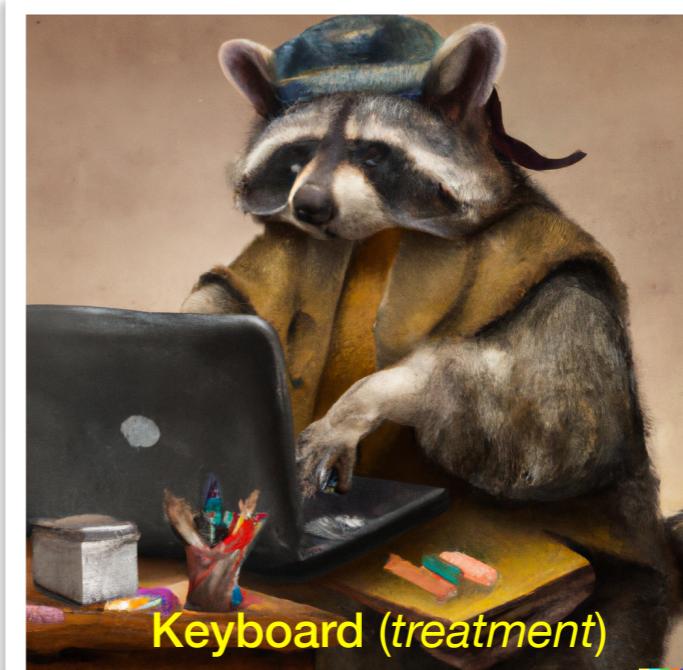
Experimental Condition: Keyboard (*treatment*)

Dependent Variables (DVs) or Effects

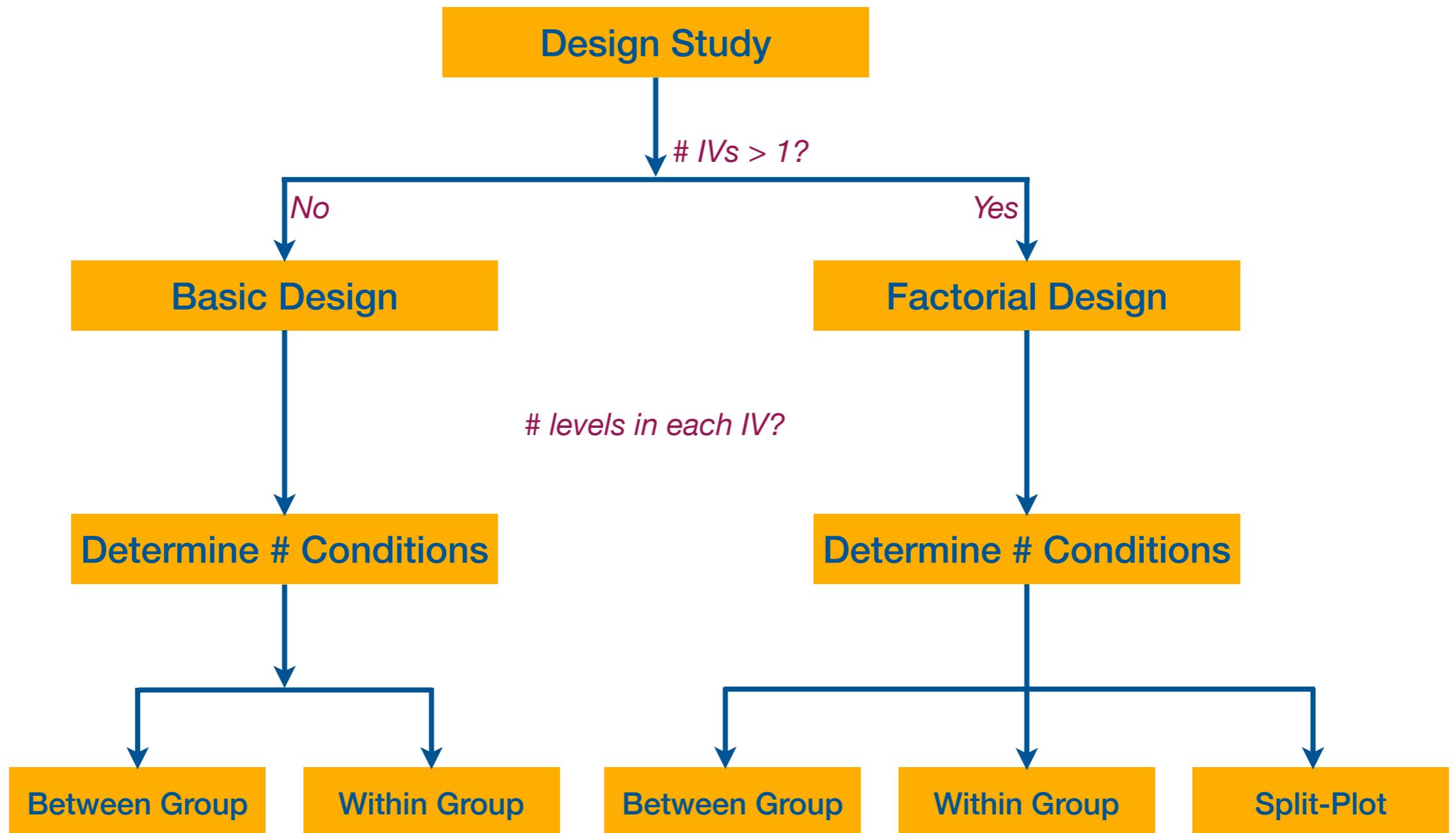
Speed

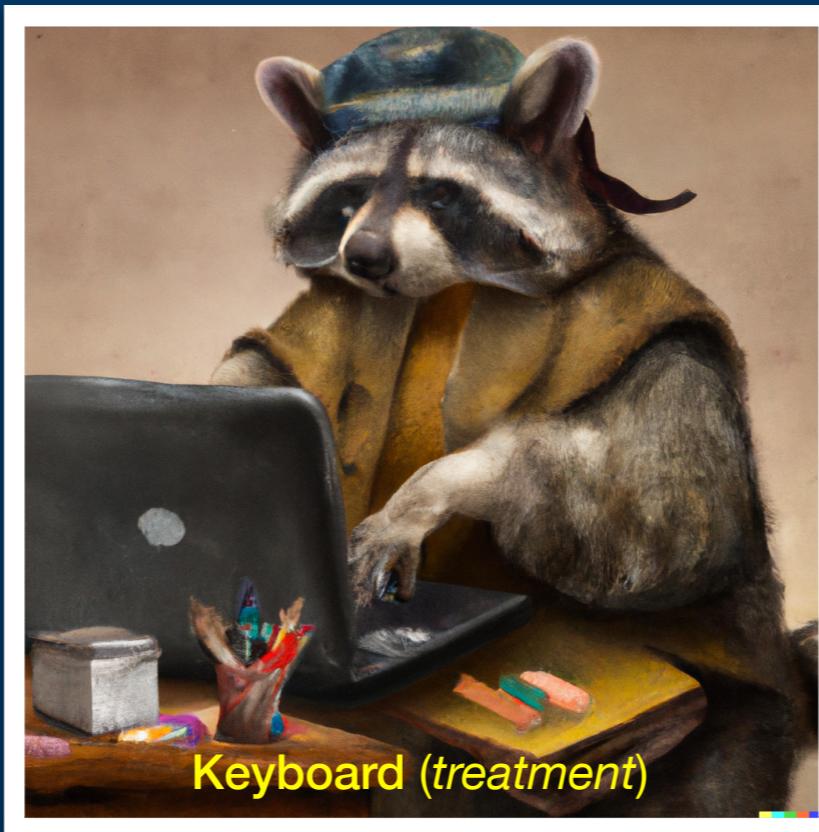
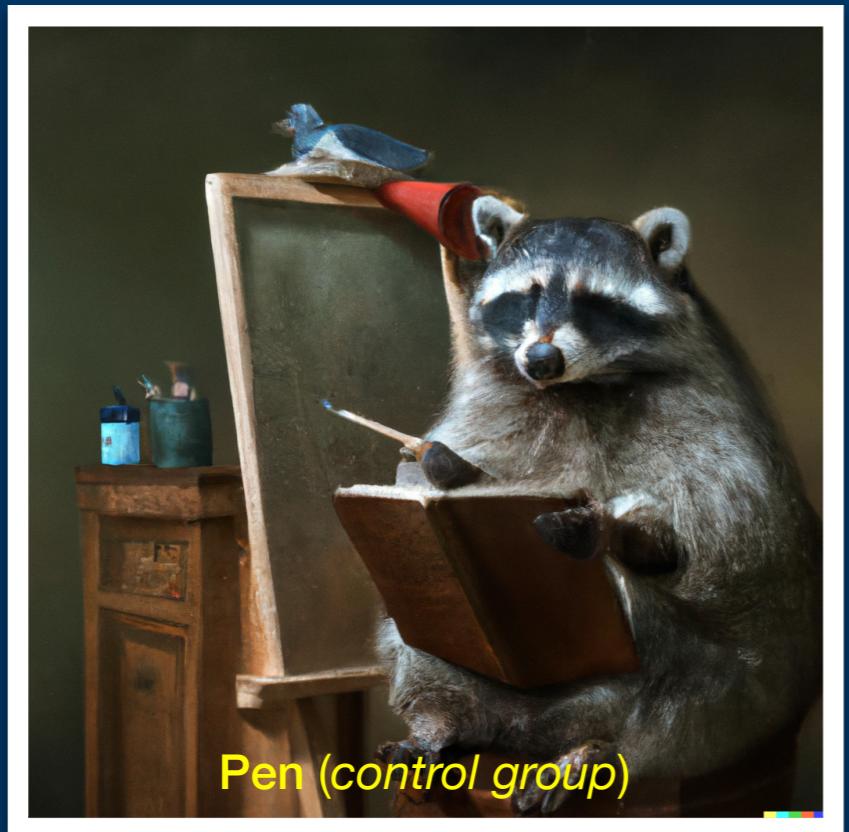
Fatigue

Usability



Experiment Design: Types



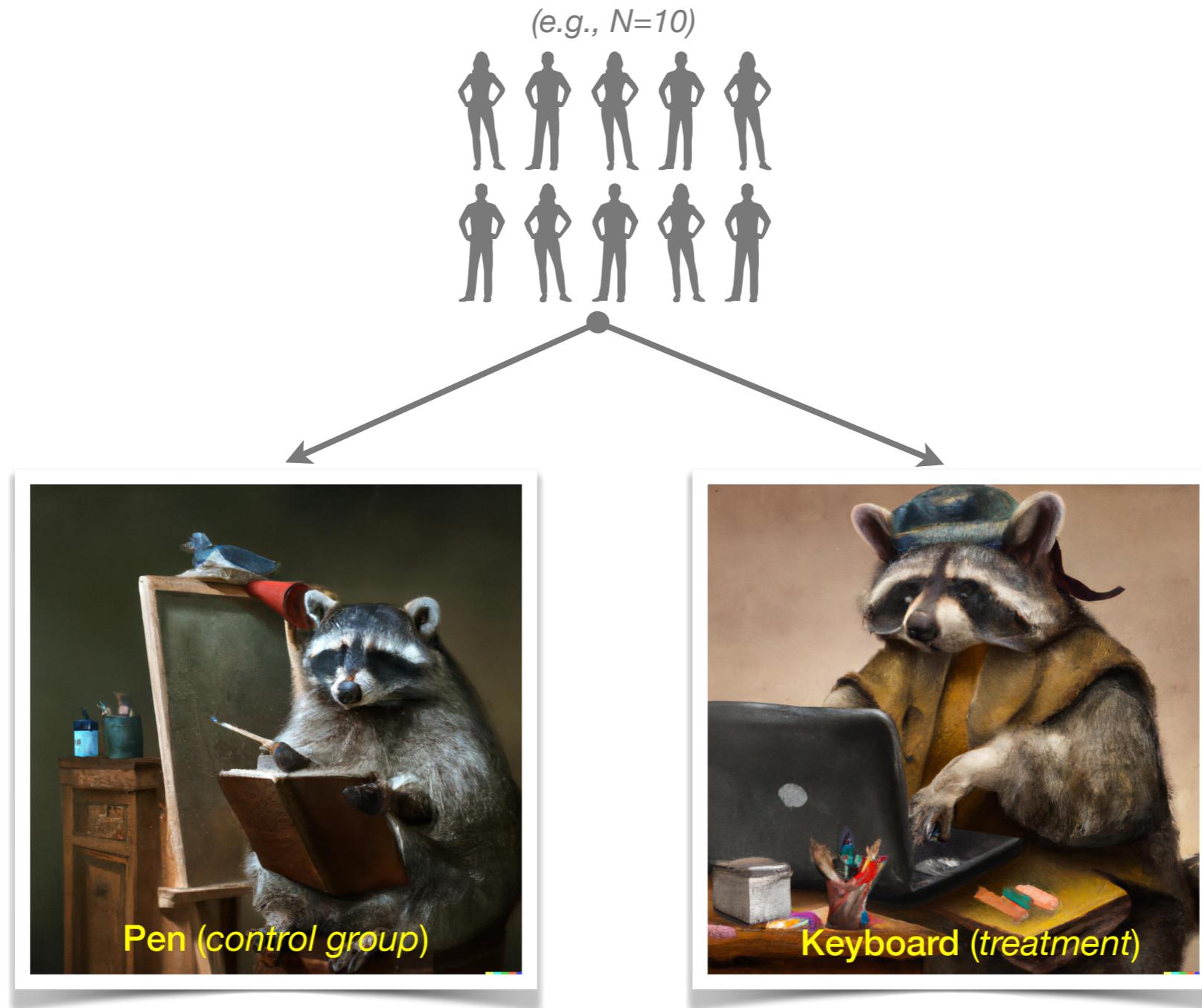


Basic Design

Basic: Between-Group Design



Basic: Within-Group Design



Between- vs. Within-Group Design

	Type of Experiment Design	
	Between-Group	Within-Group
Advantages	<ul style="list-style-type: none">- Cleaner design- Avoids learning effects- Better control of confounding factors, such as frustration	<ul style="list-style-type: none">- Smaller sample size is needed- Effective isolation of individual differences- More powerful tests
Disadvantages	<ul style="list-style-type: none">- Large sample size is needed- Individual differences may add noise (Type II errors)- Harder to get statistically significant results	<ul style="list-style-type: none">- Difficult to control learning effect- Fatigue and frustration may negatively impact results

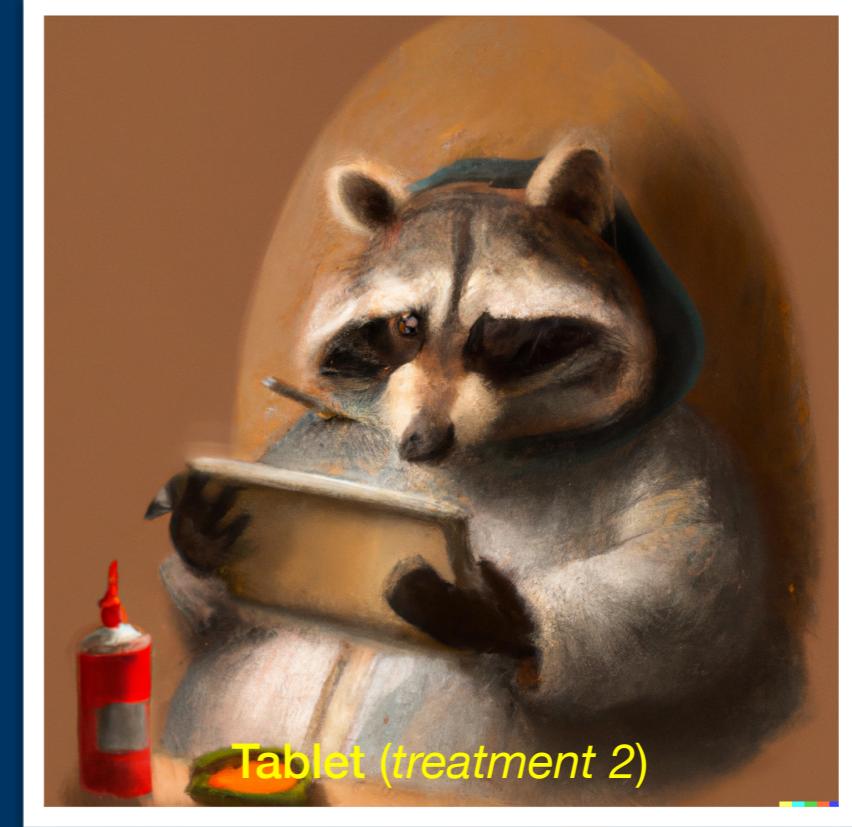
*Confounding factors are variables/aspects which influence both the IVs and DVs leading to spurious relationships. They are considered threats to internal validity.



Pen (*control group*)



Keyboard (*treatment 1*)



Tablet (*treatment 2*)

Factorial Design



Sitting



Standing



Walking

Factorial: Extra IVs

Independent Variables (IVs) or Conditions

Input Device (Pen vs. Keyboard vs. Tablet) – *3 levels*

Mobility (Sitting vs. Standing vs. Walking) – *3 levels*

Conditions: 3×3 (*9 conditions*) – “*3-by-3 design*”

Dependent Variables (DVs) or Effects

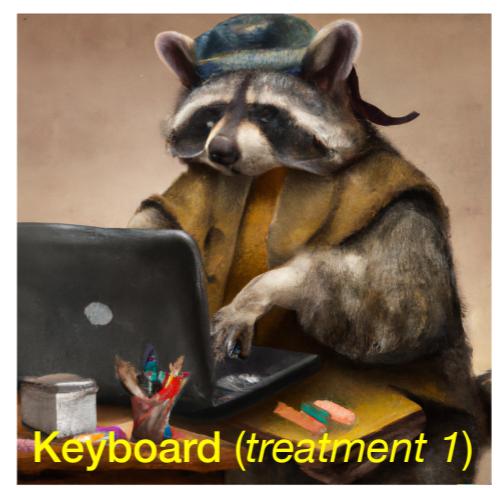
Speed

Fatigue

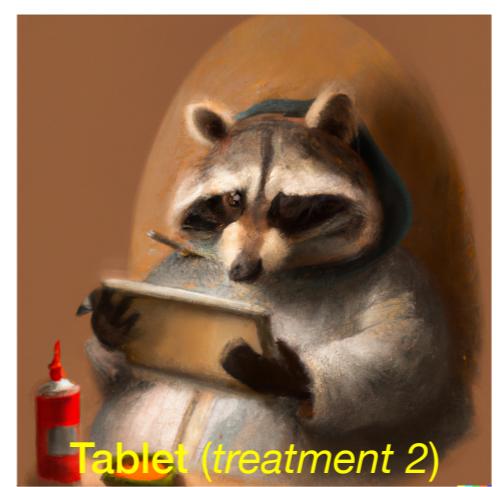
Usability



Sitting



Standing



Walking

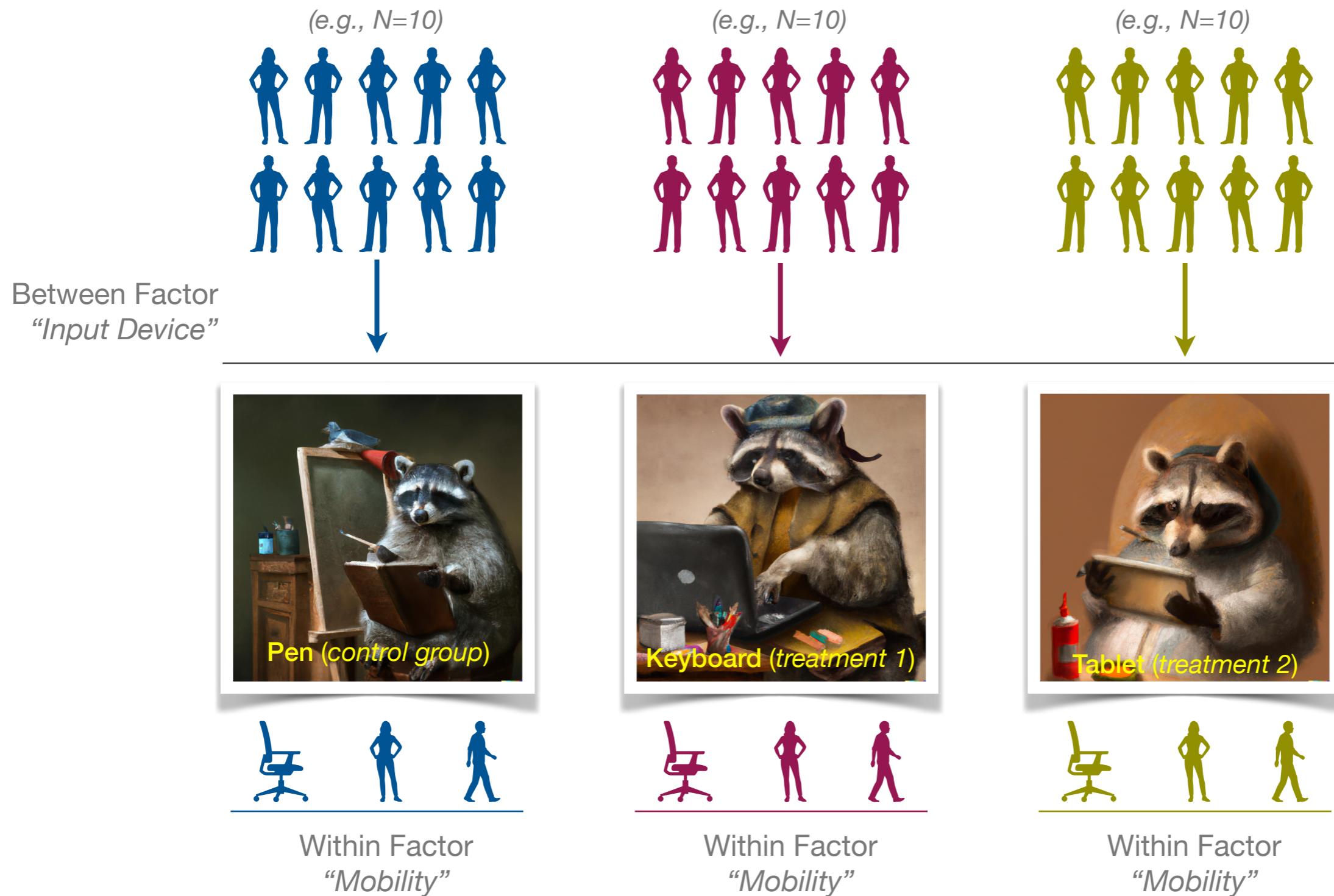
Factorial Design

		Input Device (IV)		
		Pen	Keyboard	Tablet
Mobility Type (IV)	Sitting	Speed, Fatigue, Usability	Speed, Fatigue, Usability	Speed, Fatigue, Usability
	Standing	Speed, Fatigue, Usability	Speed, Fatigue, Usability	Speed, Fatigue, Usability
	Walking	Speed, Fatigue, Usability	Speed, Fatigue, Usability	Speed, Fatigue, Usability

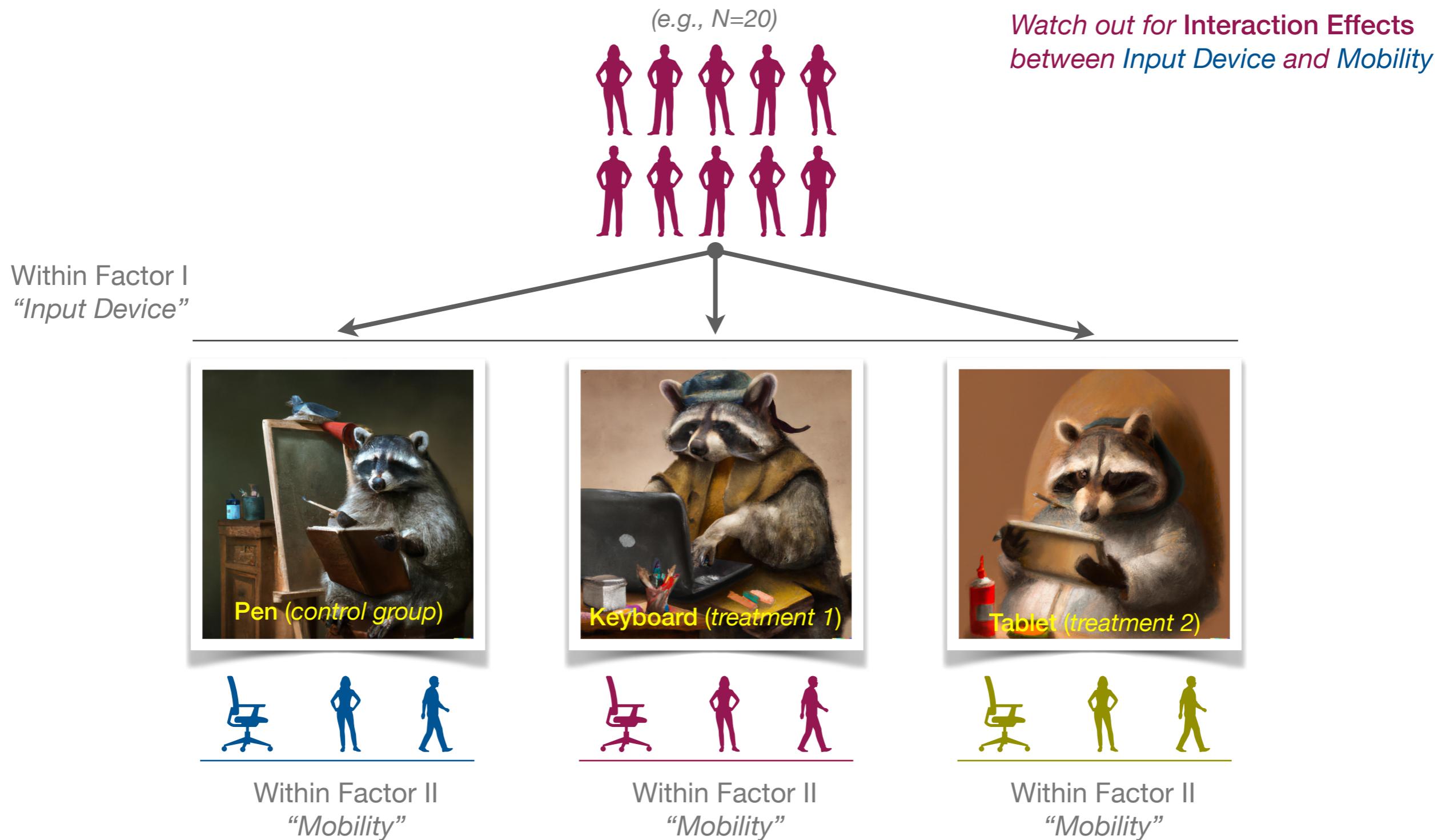
A “3-by-3” design

How would the between- and within-group work in factorial design?

Split-Plot (mixed) Design



Full Factorial Design



Errors and Biases

Errors

Random

Also called “**chance errors**” or “**noise**”, they are **bi-directional** variations around actual value under observation.

They **cannot be eliminated**, but their effect can be reduced by increasing sample size.

$$Observed = Actual \pm \epsilon$$

Systematic

Also called “**biases**”, they are **one-directional** variations which add **skewness** to the collected data.

Being **inevitable**, they truly reduce the reliability of experimental results, and should be accounted for in experiment design.

Fluctuations in experimental results, due to the study of human behavior and social interaction, is referred to as errors.

Types of Biases (Systematic Errors)

Measurement Instruments

When measurement instruments are inappropriate, inaccurate & not calibrated.

Carefully and regularly examine your instruments; use instruments which are extensively tested.

Experimental Procedures

When experimental procedures are inappropriate & unclear.

Counter-balancing tasks/treatments; clear and unambiguous instructions to participants; consistent protocols; multiple pilots to identify biases

Participants

When participants are in specific age group or have particular expertise, experience, education & personal interest.

Participants should be representative of the target user group; it's about the prototype and not the human; be mindful of their comfort and stress during experiment

Experimenter's Behavior

When experimenter (un)intentionally influence the experiment results. Intentional influence is “unethical”.

Be neutral about prototype/experiment; don't make participants wait; prepare in advance; follow strict protocol

Environmental Factors

When physical (noise, temperature, etc.) and social (# people in vicinity) impact participant's performance.

Ensure the room is quite, well-lit, comfortable, clean & prepared; if possible leave participant alone; ensure no disturbances/distractions

References

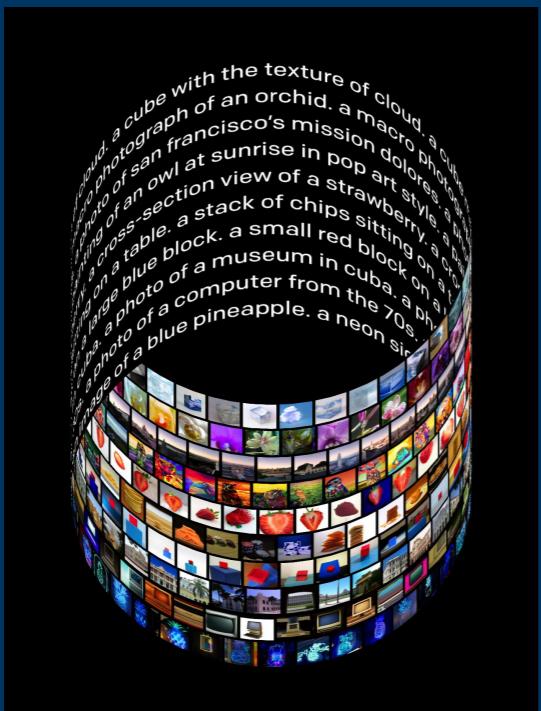
- Book: Research Methods in Human-Computer Interaction (*Jonathan Lazar, Jinjuan Feng, Harry Hochheiser*)
 - Chapter 2: Experimental Research
 - Chapter 3: Experimental Design



Acknowledgements



Dr. Evangelos Niforatos for his valuable contributions with the design of this course and for significant proportion of lecture content.



Dall-E for some of the illustrations which have been used in this course.