If evaluations are not designed properly then they do not analyze the User Interface design properly. Meaning we cannot tell if the design is good or not.

Good design is based on: understanding of how the data should be collected and how the data should be analyzed. And the related ethics. And an understanding of the scientific method.

The Scientific Method

* Obtain a hypothesis, this should be refutable
* Test hypothesis, by trying to prove it wrong.
* Analyze the outcome.

What is inductive reasoning? AKA the bottom-up approach

* Start with observations (empiricism)
* The more data we have, the truer the conclusion.
* Induction will never lead to certain knowledge because things change. For example, the sun rises on the east side. But, will that be the case tomorrow? There will be a high probability but not 100% certainty.
* Also think about the all swans are white example. The question here is have you seen all swans in the worlds, in other words you cannot say for 100% until you have the whole population.
* Thus, aims to arrive at a probability.

What is deductive reasoning? AKA the top-down approach

* Aims to arrive at a certainty.
* Start with a premise (hypothesis)
* Generate more premises
* Until a **certain** conclusion is reached.
* If a false, conclusion is deduced. Check the premises.
* Example:
* All good students pass all their exams
* Tim is a good student
* Tim will pass all his exams
* If Tim does not pass his exam, he is not a good student, or not all good students pass all their exams.



Logical Positivism: combinations of empiricism (observations) and rationalism (deduction) form the scientific method.

Forms of the scientific method:

1. If 2 instances of a phenomena have one of several causal circumstances in common, then that circumstance is the cause of the phenomena of interest. John Stuart Mill.
2. Indirect method of difference

**Internal Validity**: This is when you test on a certain population and say conclusions are valid on that population.

**External Validity**: This is when you generalize the results from the smaller populations to a wider population.

There is **no 100% certainty**. We can have **a level of confidence** as to how the findings relate to the populations.

Variables

Subject variable: Age, weight, gender of participant.

Behavioral variable: User demographics used analyses and in participant selection to create a representative population. In UX this is the user. AKA demographics or subject variables.

Stimulus variable: The interface or computer system. AKA Independent variable, input or cause.

Observable response: The thing we are measuring. AKA dependent variable, output, or effect.

Constant variable: should not change.

Extraneous variable: any variable other than the independent variable that might affect the dependent variable.

Confounding variable: factors that affect the internal and external validity. Can be controlled by preparing the lab setting. **Single or double blind procedures. Use observers. Automate the process.**

Measuring Variables

* Nominal scale: Denotes identity
* Ordinal scale: Denotes identity and magnitude
* Interval scale: Denotes identity and magnitude and has equal intervals
* Ratio scale: Has the properties of the above.
* Nominal Variable: A variable that has no numerical value.
* Ordinal Variable: Variables which can be ordered, but cannot be used to perform meaningful arithmetic like operations.
* Cardinal Variable: A variable whose values are ordered and can be multiplied by a scalar. And the magnitude between the differences in the variables is meaning. For examples wages and population.
* Interval variable: