



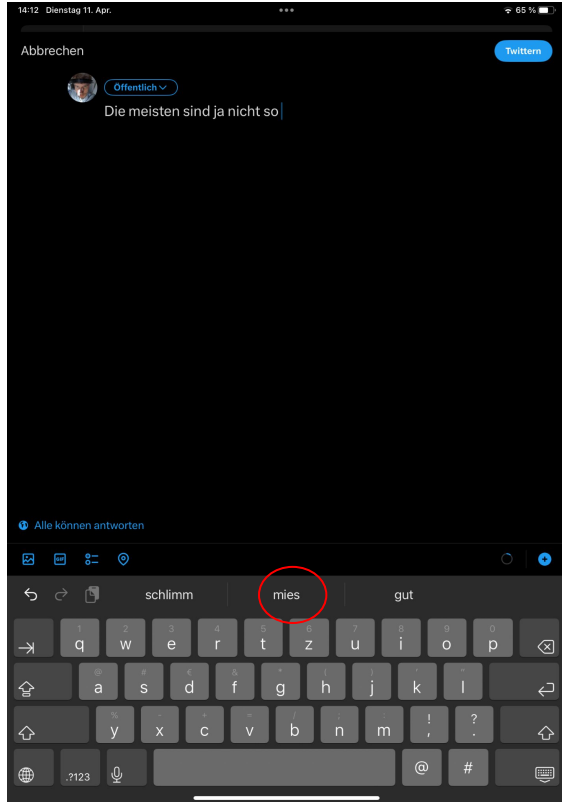
# Einführung in die Statistik

Praktische Übung – Jürgen Hermes – IDH – SoSe 2023

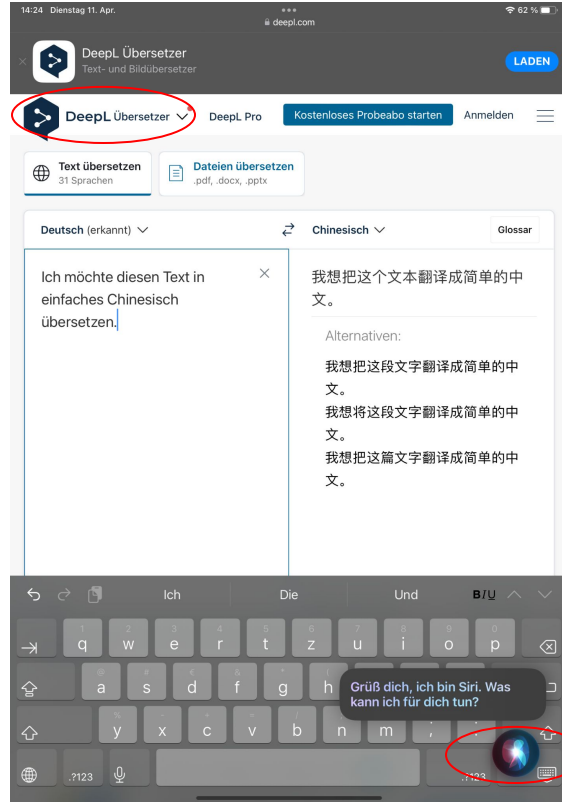
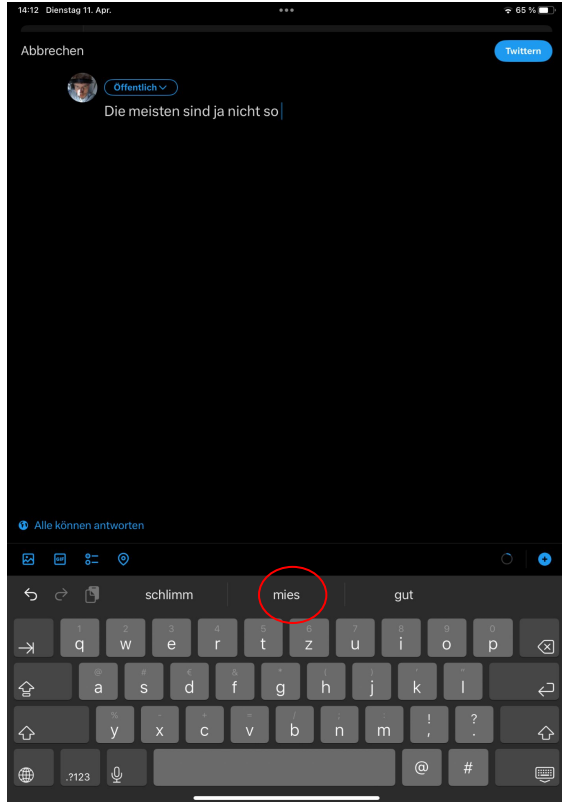
# Wofür benötigen wir Statistik und Wahrscheinlichkeitstheorie?

- Statistik und Wahrscheinlichkeitstheorie sind zusammengefasst im mathematischen Gebiet **Stochastik**, in dem es um die *mathematische Modellierung zufälliger Ereignisse* geht.
- Solide stochastische Grundkenntnisse sind unentbehrliche Grundlage für die Durchführung in die Bewertung von **empirischer Forschung**.
- Stochastische Methoden sind außerdem Grundlage der gegenwärtig (und wohl auch zukünftig) erfolgreichsten Ansätze der Sprachverarbeitung und der Künstlichen Intelligenz (Stichwort **Sprachmodelle**).

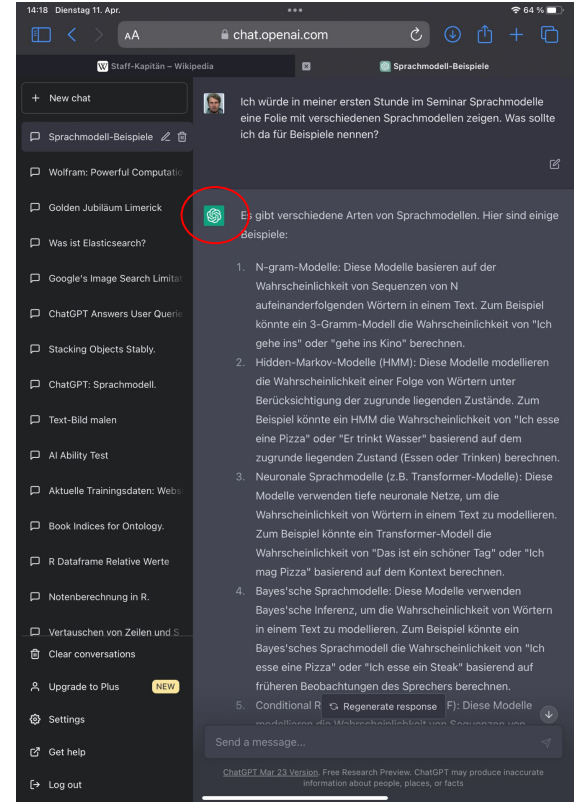
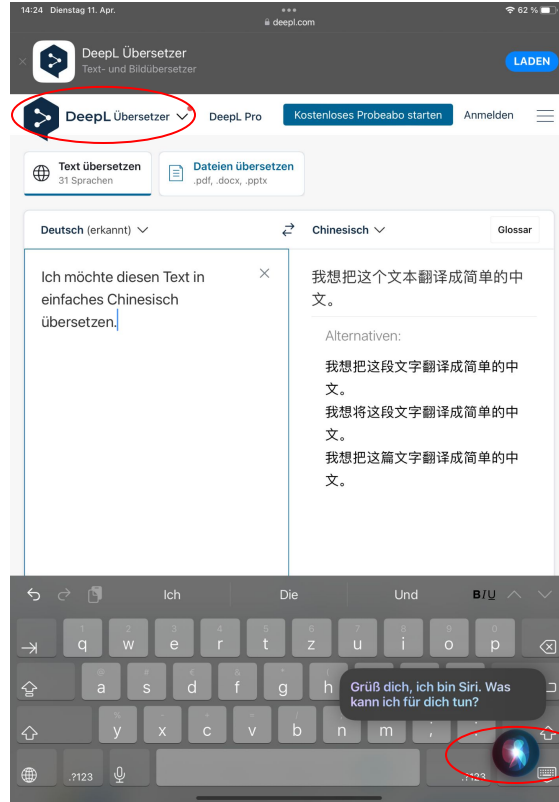
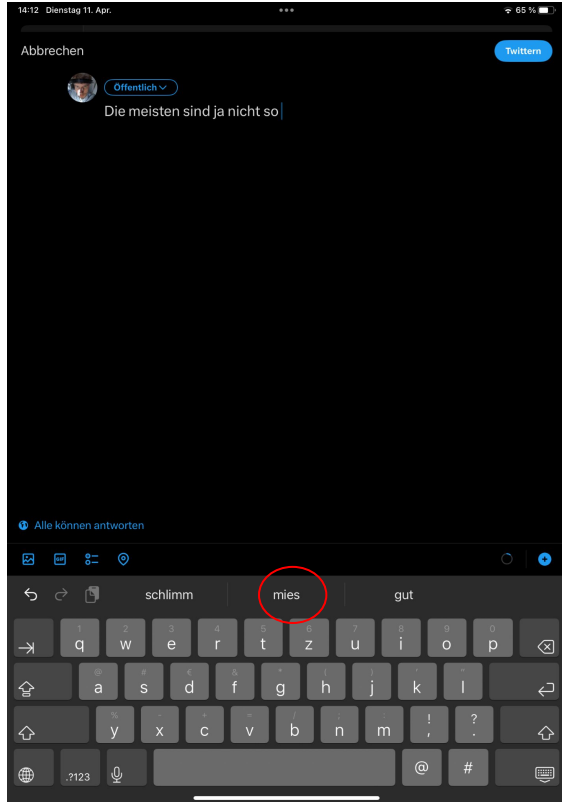
# Beispiele für aktuell verwendete Sprachmodelle



# Beispiele für aktuell verwendete Sprachmodelle



# Beispiele für aktuell verwendete Sprachmodelle





Diese und die folgenden Folien sind erstellt worden von Sascha Wolfer für seinen Kurs “Statistik mit R” an der Uni Basel. Ich nutze sie mit seiner freundlichen Genehmigung. DOI für die Materialien ist

[10.5281/zenodo.7431504](https://doi.org/10.5281/zenodo.7431504)

# Prolog: Grundbegriffe empirischer Forschung





Was würden Sie sagen/denken,  
wenn ich Ihnen die folgende Aufgabe  
gäbe?

Gehen Sie auf den  
Marktplatz und  
beobachten Sie dort die  
Menschen!

- Was soll ich beobachten?
- Worauf soll ich achten?
- Was will er denn überhaupt wissen?

„Ohne Problem keine Beobachtung.

- Karl Popper

# Fragestellung, Theorie, Hypothese, Beobachtung

- Aus einer **Fragestellung** entsteht eine empirische **Theorie**.
  - Theorie: System logisch konsistenter Aussagen, das an dem Ausschnitt der Realität, den die Theorie zu erklären versucht, scheitern kann.
  - Aus der Theorie abgeleitete **Hypothesen** müssen sich also an der Welt messen lassen. Eine Theorie ohne Bezug zur Welt ist bedeutungslos.
- **Beobachtungen** verbinden Theorien mit der Welt.



# Fragestellung, Theorie, Hypothese, Beobachtung

## Fragestellung

Welchen Effekt hat eine Impfung gegen Covid-19 auf die Schwere der Erkrankungsverläufe?

## Theorie

Impfungen gegen Krankheiten verringern die Gefahr einer schweren Erkrankung.

## Hypothese

Geimpfte Menschen haben ein geringeres Risiko an einer Covid-19-Erkrankung zu versterben.

## Beobachtung

Vergleichende Studie zwischen geimpften und ungeimpften Personen.



# Hypothesen

- Eine Hypothese im empirischen Sinn ist eine Aussage in Form einer **überprüfbaren** Behauptung.
- Wissenschaftliche Hypothesen nach Bortz & Döring (2016):
  - **Allgemeingültig**, über den Einzelfall oder ein singuläres Ereignis hinausgehende Behauptung
  - Kann in einen sinnvollen **Konditionalsatz** überführt werden ("wenn ... dann", "je ... desto ...")
  - Muss potentiell **falsifizierbar** sein. Es müssen Ereignisse denkbar sein, die dem Konditionalsatz widersprechen.

# Hypothesen

## gerichtete Hypothese

Rauchen kann gesundheitsschädlich sein.

Menschen, die rauchen, haben mehr schwere Krankheiten als Menschen, die nicht rauchen.

## ungerichtete Hypothese

Bestimmte Wörter werden in Wörterbüchern deutlich häufiger nachgeschlagen als andere.

Die Nachschlagehäufigkeit eines Worts hängt mit der Auftretensfrequenz des Wort in der Sprache zusammen.

# Nullhypothese und Alternativhypothese



**Nullhypothese:** Es gibt *keinen* Zusammenhang oder Unterschied (= Effekt).



**Alternativhypothese:** Es gibt einen Effekt.

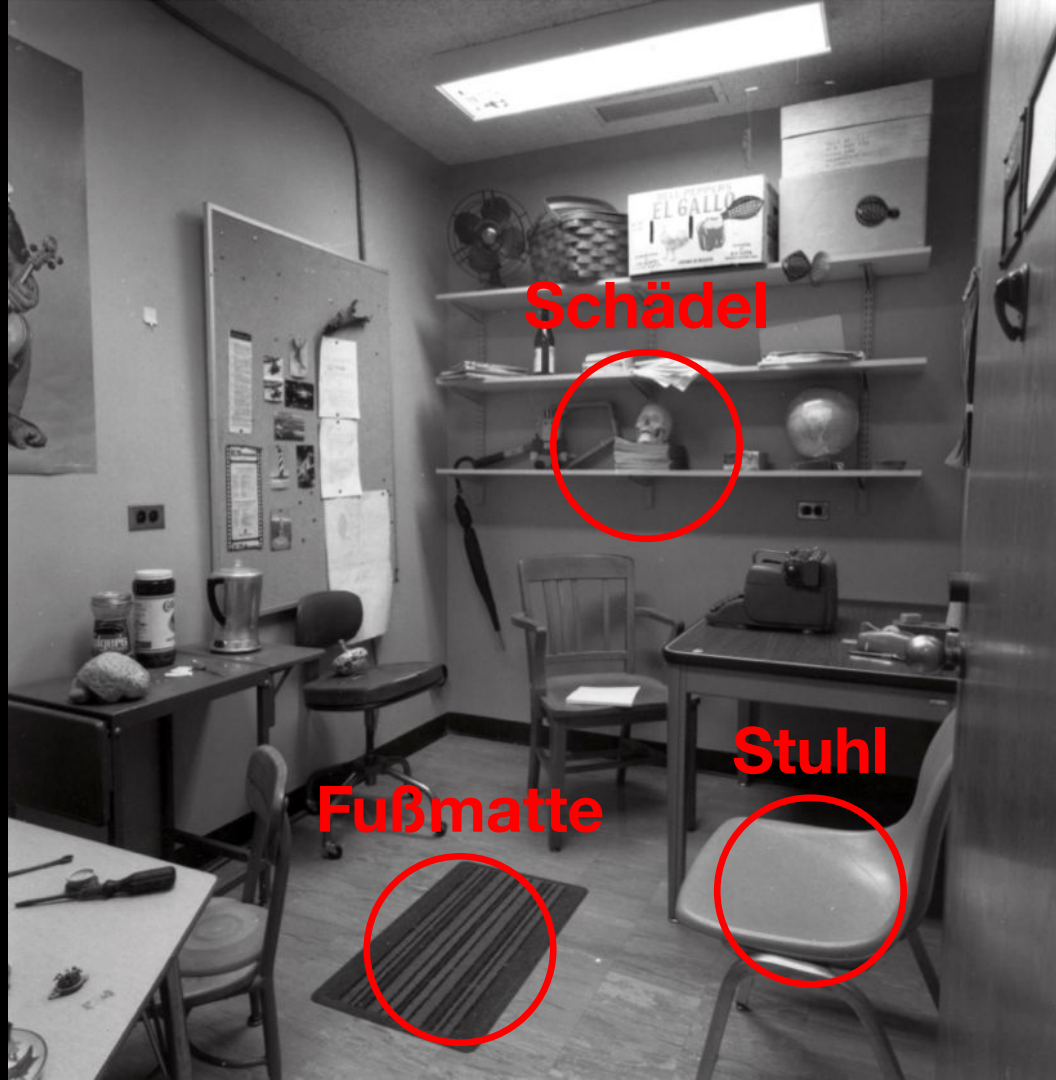
- Die statistischen Test, die Sie hier lernen, prüfen immer die **Nullhypothese**.
- Wenn wir diese mit ausreichender Wahrscheinlichkeit ablehnen können, nehmen wir an, dass die **Alternativhypothese** gilt.
  - Alternativhypothese = Forschungshypothese.



# Experiment

- Was haben Sie gesehen?
- Haben Sie einen Stuhl gesehen?
- Haben Sie ein Buch gesehen?
- Haben Sie einen Schädel gesehen?
- Haben Sie einen Computer gesehen?
- Haben Sie eine Fußmatte gesehen?





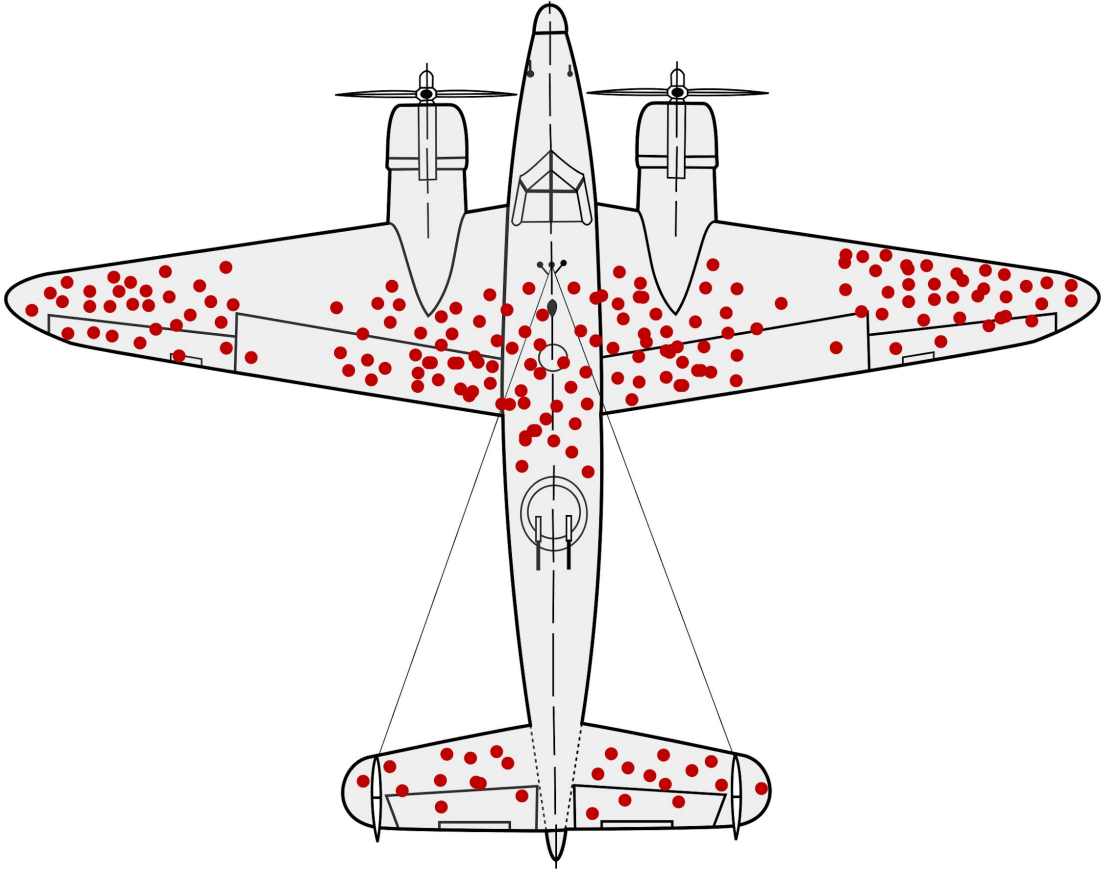
Schädel

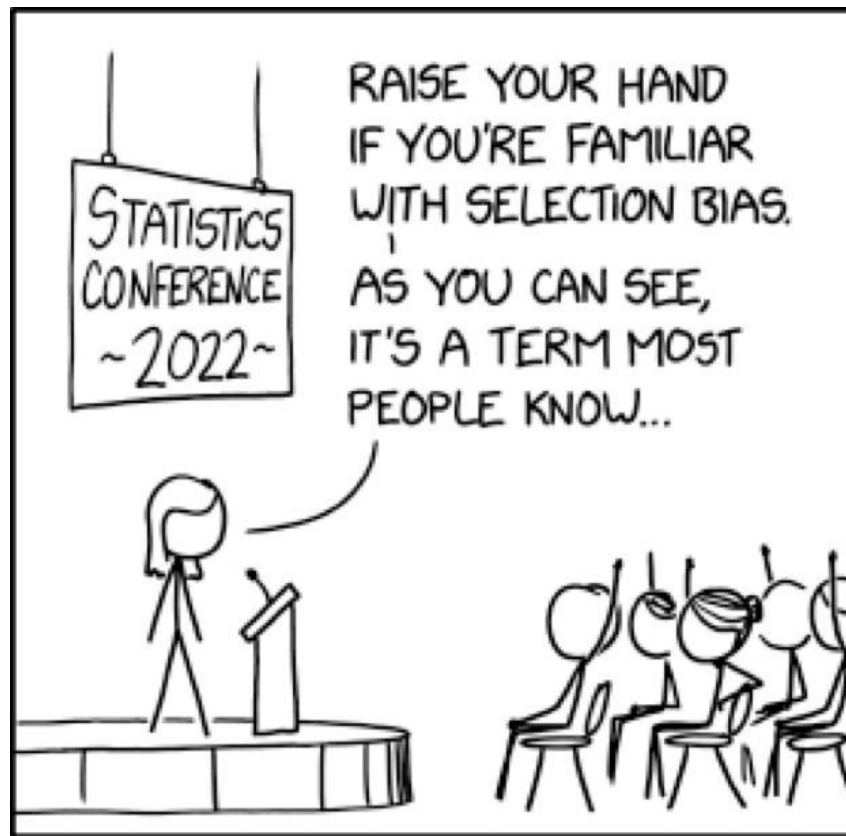
Fußmatte

Stuhl

# Verlässlichkeit von Beobachtungen

- Das Problem: Beobachtungen sind nicht immer verlässlich.
  - Ungenaue oder unangemessene Messinstrumente
  - Erinnerung rekonstruktiv und fehlerbehaftet
- Was bedeutet eine bestimmte Beobachtung?
- Welche Beobachtungen sind relevant?
  - Dominanz konkreter Ereignisse über abstrakte Grundraten
    - “Anekdotische Evidenz”, Veränderungsblindheit
- *Biases* (z.B. *confirmation bias*, *survivorship bias*, *selection bias*)





# Gütekriterien empirischer Forschung

## Objektivität

Die Testergebnisse sind unabhängig von der Person, die den Test durchführt.

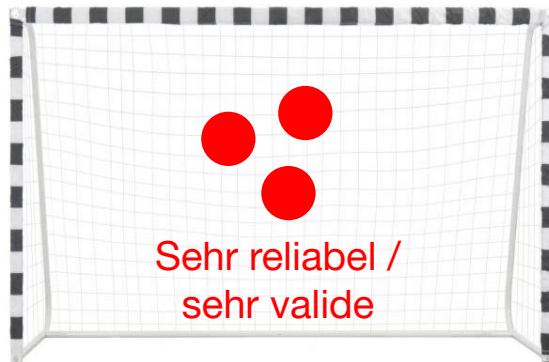


Nicht reliabel /  
nicht valide



## Reliabilität

Es wird möglichst genau gemessen.



## Validität

Der Test misst das, was er zu messen vorgibt.



?

Sehr reliabel /  
nicht valide

# Objektivität

- **Durchführungsobjektivität:** Ergebnisse sind unabhängig von der forschenden Person.
  - Gegenbeispiel: Männer geben systematisch niedrigere Werte des Schmerzempfindens gegenüber weiblichen Testleiterinnen an.
- **Auswertungsobjektivität:** Gleiches Verhalten der Testpersonen wird gleich ausgewertet.
  - Nicht gegeben: Zwei Dozierende bewerten dieselbe Klausur. Dozent:in A benutzt vorgefertigte Stichwortlisten; Dozent:in B sucht nach inhaltlichen Argumentationsfehlern.
- **Interpretationsobjektivität:** Gleiche Ergebnisse werden gleich interpretiert.
  - 15 Punkte in einer Klausur bedeutet immer "sehr gut". Unter 5 Punkte bedeutet immer "nicht bestanden".

# Reliabilität

- **Retest**-Reliabilität: Unter identischen Bedingungen muss bei wiederholten Messungen das gleiche Ergebnis herauskommen.
- **Interrater**-Reliabilität: Mehrere Personen müssen zum gleichen Ergebnis kommen (bspw. bei Annotationen oder Einschätzungen).
- **Split-half**-Reliabilität: Wenn ich den Datensatz in zwei Hälften teile, sollten die statistischen Kennwerte in beiden Hälften ähnlich sein.

# Validität

- **Interne Validität:** Wie gut misst der Test unter Laborbedingungen das, was er messen soll?
  - Laborbedingungen: konstant und kontrolliert
- **Externe** (ökologische) Validität: Sind die Ergebnisse des Tests auf die "wahre Welt" übertragbar?
  - Typischerweise: *Trade-off* zwischen externer und interner Validität
- **Konstruktvalidität:** Wie gut ist das zu messende Merkmal operationalisiert (= messbar gemacht)?



# Operationalisierung

- Wie messen wir die Variable, die uns interessiert?
  - Ergebnis sollte ein konkretes Messergebnis sein (Zahl, Kodierung).
- Recherche notwendig:
  - Wie haben andere Forschungsteams dieselbe Variable gemessen?
  - Standardisierte Verfahren?
- Oft mehrere Operationalisierungen denkbar
- Viele wissenschaftliche Diskussionen drehen sich um die geeignete Form der Operationalisierung.

# Begriffe

**Fragestellung**

**Nullhypothese**

**Konstruktvalidität**

**Theorie**

**Alternativhypothese**

**Operationalisierung**

**Hypothese**

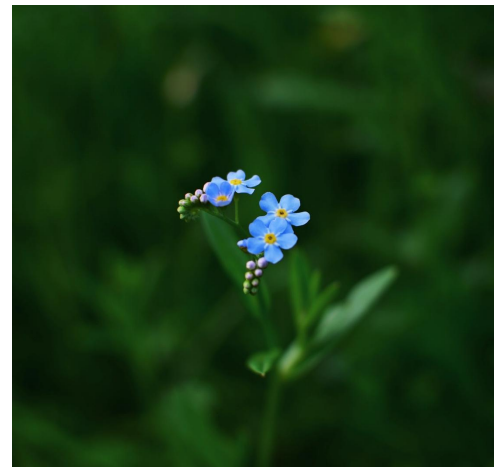
**Objektivität**

**Falsifizierbarkeit**

**Reliabilität**

**gerichtete /  
ungerichtete  
Hypothese**

**interne / externe  
Validität**



# Die Programmiersprache R

- Freie, vollständige Programmiersprache, die meist für statistische Berechnungen und Visualisierungen eingesetzt wird.
- Publiziert 1993, aktuelle Version 4.1.3
- Benutzerfreundliche integrierte Entwicklungsumgebung RStudio
- Besonderheiten: Für statistische Berechnungen optimierte Datenstrukturen und Funktionen, besonderes Potential in der einfachen Grafikerzeugung.
- “To understand computations in R, two slogans are helpful: Everything that exists is an object. Everything that happens is a function call.” (John Chambers)
- Extrem große Nutzer:innen-Community, die viele vorgefertigte, nutzbare Pakete zur Verfügung stellt.
- Extrem große Verbreitung innerhalb der Forschungsrichtung Data Science → Das Erlernen von R hilft bei einer Vielzahl potentieller Berufe.

# Hausaufgabe

- Falls noch nicht geschehen: Bearbeiten Sie die Hausaufgabe aus der letzten Woche (Installation von R und RStudio)
- (optional) Wenn Sie etwas mehr Kontext benötigen, lesen Sie das erste Kapitel “Why is my evil lecturer forcing me to learn statistics?” aus dem Buch “Discovering Statistics using R” (Dieses finden Sie als “sample chapter” hier: <https://www.discoveringstatistics.com/books/discovering-statistics-using-r/>) Sie können das Buch auch aus der Lehrbuchsammlung ausleihen, es geht mit seinen mehr als 1000 Seiten aber weit über unseren Kurs hinaus, vielleicht haben Sie aber ja gerne eine Art von Nachschlagewerk.