

# *Forschungsparadigmen und Methoden für RSEs*

## *Sitzung 4: Grundlagen quantitativer Forschung für RSE*



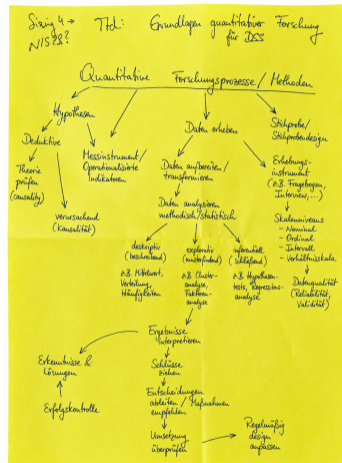


Figure 2: ChatGPT Version 1

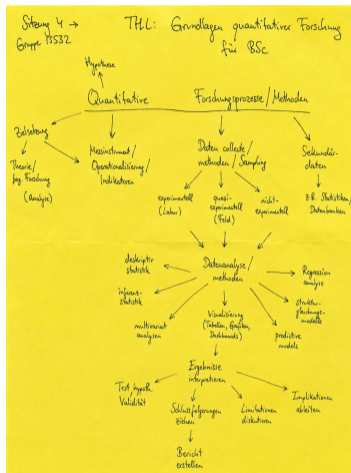


Figure 3: ChatGPT Version 2

## *Das quantitative Experiment (Zapf 2013)*

„Wiederholbare Beobachtungen unter kontrollierten Bedingungen, wobei eine (oder mehrere) unabhängige Variable(n) derartig manipuliert werden, dass eine Überprüfungsmöglichkeit der zugrunde liegenden Hypothese (Behauptung eines Kausalzusammenhangs) in unterschiedlichen Situationen gegeben ist.“

- ▶ Experiment funktioniert nur, wenn der Reiz zu Realverhalten führt
- ▶ Exakteste Messmethode, aber nicht immer durchführbar
- ▶ Untersuchung der Wirkung wesentlicher Variablen
- ▶ Misst objektive, subjektive, materielle und ideelle Determinanten
- ▶ Analysiert numerische Daten über Beziehungen von Verhaltensvariablen
- ▶ In kontrollierten Situationen
- ▶ Vereinfacht komplexe Zusammenhänge
- ▶ Testet Hypothesen
- ▶ Unabhängige Variable wird manipuliert

→ Verhaltensdifferenzen als Resultat der Manipulation werden gemessen

## *Vor- und Nachteile (Zapf 2013)*

### **Vorteil:**

- ▶ Exakteste Messung (bei isolierten, kontrollierten, manipulierten Variablen)

### **Nachteil:**

- ▶ Kenntnis wesentlicher Variablen nötig
- ▶ Einfluss von Störvariablen muss vermieden werden
- ▶ Hoher Schulungsaufwand

## *Feldexperiment:*

- ▶ Authentische Situation
- ▶ Untersuchungsobjekt bleibt in natürlicher Umgebung
- ▶ Zwei kontrastierende Gruppen
- ▶ Nicht alle Variablen kontrollierbar

## *Laborexperiment:*

- ▶ Künstliche, kontrollierte Situation
- ▶ „Reine“, vereinfachte Bedingungen
- ▶ Experimentalgruppe + Kontrollgruppe
- ▶ Alle Variablen kontrollierbar
- ▶ Fokus: Wirkung der unabhängigen Variable

### *Ex-post-facto Experiment (Quasiexperiment):*

- ▶ Untersuchung vergangener sozialer Vorgänge
- ▶ Rückverfolgung zu kausalen Faktoren
- ▶ Nutzung von Dokumenten
- ▶ Beispiel: Einfluss des Abiturs auf sozialen Erfolg

## *Kontrollgruppenexperiment:*

- ▶ Mindestens zwei ähnliche Gruppen (Zufallsverteilung)
- ▶ Eine Gruppe wird manipuliert, eine nicht
- ▶ Vergleich von Vorher-/Nachher-Messungen
- ▶ Verhaltensdifferenz wird gemessen
- ▶ Nachteil: Gruppen nie exakt identisch

## *Zeitreihenexperiment:*

- ▶ Gleiche Gruppe mehrfach gemessen
- ▶ Sequenz: ohne – mit – ohne Einwirkung
- ▶ Verhaltensdifferenz wird gemessen
- ▶ Nachteil: Gruppenveränderung (z. B. Ausstieg)

### *Voraussetzungen (Zapf 2013)*

1. Variablen und Bedingungen müssen kontrollierbar sein
2. Hypothesen müssen Zusammenhänge zwischen Variablen formulieren
3. Verhaltensvariablen müssen isolierbar sein
4. Unabhängige Variable muss manipulierbar sein
5. Manipulation muss wiederholbar sein

### *Anwendungsgebiete (Zapf 2013)*

1. Maximierung und Minimierung
2. Testen von Grenzen
3. Adaption (Anpassung)

## *Einwände (Zapf 2013)*

1. Self-fulfilling-prophecy (Verhalten im Experiment anders als in Realität)
2. Self-destroying process (Aufgabe wird so stark internalisiert, dass Verhalten verzerrt wird)
3. Ethische (moralische) Einwände (Grenzen individuell festzulegen, kein dauerhafter Schaden)
4. Selektivität des Experiments (Erklärt nicht das „wahre Leben“)

## *Arbeitsschritte (Zapf 2013)*

- ▶ Festlegung der bedeutsamen Faktoren
- ▶ Hypothesen aufstellen
- ▶ Manipulierte Variable festlegen
- ▶ Bestimmung von Experimental- und Vergleichsgruppe (nach soziodemographischen Faktoren oder Zufallsauswahl)
- ▶ Konstruktion der Experimentalsituation (wer, wann, wo, wie, womit)
- ▶ Prätest

Inhaltsanalyse meint das Interpretieren von Symbolen mit wissenschaftlichen Methoden.

## *Symbol (Zapf 2013)*

- ▶ Marker für Begriffe
- ▶ Menschliches Denken ist an Symbolgebrauch gebunden
- ▶ Symbole haben Bedeutung für Menschen
- ▶ Erschließt sich die Bedeutung → löst Handlungen aus
- ▶ Voraussetzung für Interaktionen

## *Merkmale (Zapf 2013)*

- ▶ Hat Kommunikation zum Gegenstand
- ▶ Arbeitet mit symbolischem Material (fixierte Kommunikation)
- ▶ Geht systematisch vor
- ▶ Ist regelgeleitet
- ▶ Ist theoriegeleitet
- ▶ Ist eine schlussfolgernde Methode (Rückschlüsse auf Kommunikation)
- ▶ Rekonstruiert die Verteilung von Merkmalen in Populationen
- ▶ Sucht Typen von Kommunikationsmustern und -ergebnissen
- ▶ Hilfsmittel zur Überprüfung qualitativ gewonnener Hypothesen

## *Methoden (Zapf 2013)*

- ▶ Vom Analysenmaterial wird auf den Kontext geschlossen
- ▶ Hoher Systematisierungsgrad wird angestrebt
- ▶ Reduktion von Informationen
- ▶ Zuordnung von Textteilen zu Merkmalen
- ▶ Übersetzung der Objektsprache in Metasprache (sozialwissenschaftliche Kategorien)
- ▶ Diktionär = inhaltsanalytisches Wörterbuch

## *Anwendungsgebiete (Zapf 2013)*

- ▶ Informationsbeschreibung und Normenprüfung
- ▶ Informationsraffung und Informationsbeschleunigung
- ▶ Informationsgenese (Entwicklung von Inhalten in Abhängigkeit zur sozialen Wirklichkeit)
- ▶ Prognosen



<https://xkcd.com/1860/>, CC2.5

## Annahmen:

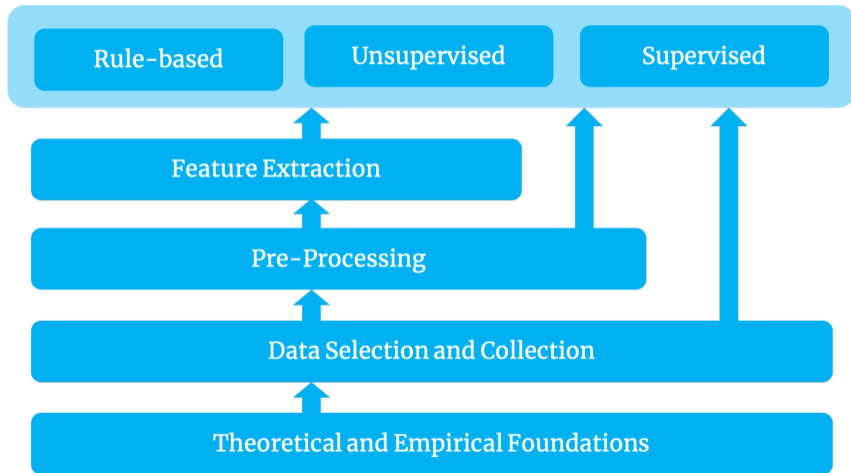
- ▶ Texte sind eine wertvolle Ressource für die Analyse
- ▶ Die interessierenden Einheiten können aus Texten extrahiert werden, z. B. durch das Messen von Merkmalen
- ▶ Bag-of-Words-Annahme
- ▶ Berücksichtigung von Kontext
- ▶ Zuverlässige Schätzungen der manifesten und/oder latenten interessierenden Konstrukte (= Inferenz)

## *Relevanz von Text als Daten (Gold 2022)*

- ▶ Texte vermitteln von Natur aus Informationen und sind daher wertvolle Quellen für systematische Analysen (Benoit 2020, S. 463).
- ▶ Die fortschreitende Digitalisierung der Welt hat zu einer stetig wachsenden Verfügbarkeit großer Textmengen geführt (Bubenhofer and Scharloth 2015, S. 2).
- ▶ Die Entwicklung und Anwendung neuer computergestützter Methoden zur Textverarbeitung, einschließlich des interdisziplinären Transfers („Migration“) von Algorithmen sowie jüngster Fortschritte bei großen Sprachmodellen wie ChatGPT.
- ▶ Die „Emanzipation der Daten“ (Bubenhofer and Scharloth 2015, S. 2): Viele Textdaten werden ursprünglich nicht zu Analysezwecken erzeugt, sondern für die Forschung zweckentfremdet.

1. Sozialwissenschaftliche Theorien und fachliches Wissen sind essenziell für das Forschungsdesign
2. Textanalyse ersetzt den Menschen nicht — sie ergänzt ihn
3. Der Aufbau, die Verfeinerung und das Testen sozialwissenschaftlicher Theorien erfordern Iteration und Kumulation
4. Textanalysemethoden destillieren Verallgemeinerungen aus Sprache
5. Die beste Methode hängt von der jeweiligen Aufgabe ab
6. Validierung ist essenziell und hängt von Theorie und Aufgabe ab

# Vorgehen (Gold 2022)



- ▶ nominalskalierte Variablen: Bezeichnungen sind eindeutig, erlaubt Aussagen über Verschiedenheit der Merkmalsträger
- ▶ ordinalskalierte Variablen: Aussage über Reihenfolge
- ▶ intervallskalierte Variablen: Differenzen zwischen Werten sind eindeutig
- ▶ verhältnisskalierte Variablen: Wie Intervallskalen aber mit eindeutigem Nullpunkt
- ▶ absolutskalierte Variablen: z.B. Häufigkeit

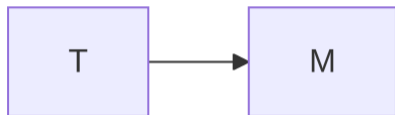
Welche der folgenden Transformationen sind (a) bei einer Nominalskala und (b) bei einer Verhältnisskala zulässig?

1. eineindeutige Transformation
2. streng monotone Transformation
3. positiv lineare Transformation
4. Ähnlichkeitstransformation
5. Identitätstransformation

## Mathematische Eigenschaften (Bortz and Döring 2006)

Skala	Alle eindeutigen Transformationen	Alle streng monotonen Transformationen	Alle positiven linearen Transformationen	Alle Ähnlichkeits-transformationen	Alle Identitäts-transformationen.
Nominal	ja	ja	ja	ja	ja
Ordinal	nein	ja	ja	ja	ja
Intervall	nein	nein	ja	ja	ja
Verhältnis	nein	nein	nein	ja	ja
Absolut	nein	nein	nein	nein	ja

## Beispiel: Bewerbungstraining für Arbeitslose



T: Treatment (Bewerbungstraining)

M: Messung (z.B. Qualität der Bewerbungen, Einladungen zum Vorstellungsgespräch)

**Störvariablen:** - Unterschiede in M können auch schon vor dem Treatment bestehen

# „Ein-Gruppen-Pretest-Posttest-Design“ (Maaß 2014)



T: Treatment (Bewerbungstraining)

M1: Pretestmessung

M2: Posttestmessung

## **Störvariablen:**

- ▶ zwischenzeitliche Einflüsse unabhängig vom Treatment
- ▶ Untersuchungsteilnehmer entwickeln sich unabhängig vom Treatment weiter
- ▶ allein die Pretestmessung verändert das Verhalten
- ▶ gemessenes Verhalten ohnehin starker Variabilität unterliegt

## *Ex-post-facto-Plan (ohne Randomisierung) (Maaz 2014)*

Vergleichsgruppe



Experimentalgruppe



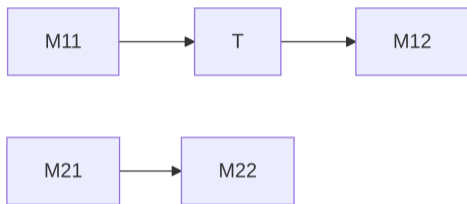
T: Treatment (Bewerbungstraining)

M1: Posttestmessung

M2: Messung Vergleichsgruppe

**Störvariablen:** Unterschiede zwischen Gruppen können schon vorher bestehen

# Kontrollgruppenplan mit Pre- und Posttest (ohne Randomisierung) (Maaz 2014)



T: Treatment

(Bewerbungstraining)

M11 und M12: Pre- und  
Posttestmessung

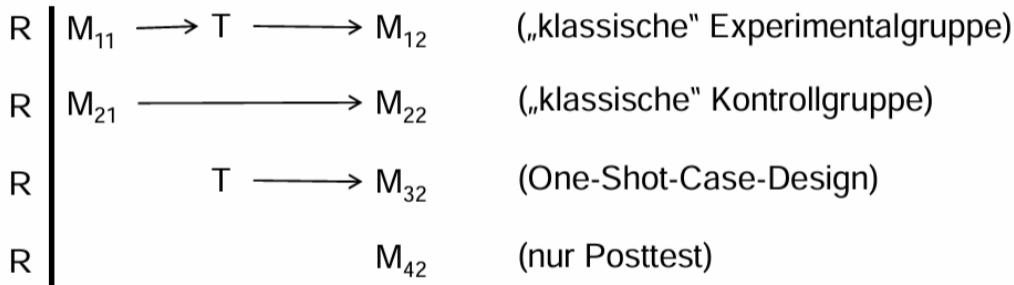
Experimentalgruppe

M21 und M22: Pre- und  
Posttestmessung Kontrollgruppe

## Störvariablen:

- ▶ Experimental- und Kontrollgruppe können sich in nicht-beobachteten Variablen unterscheiden (z.B. Motivation eine Stelle zu finden)
- ▶ unerwünschte Effekte der Pretestmessung

## Solomon-Vier-Gruppen-Plan (Maaz 2014)



- ▶ „klassische“ Experimentalgruppe:  $M11 \rightarrow T \rightarrow M12$
- ▶ „klassische“ Kontrollgruppe:  $M21 \rightarrow M22$
- ▶ One-Shot-Case-Design:  $T \rightarrow M32$
- ▶ nur Posttest:  $M42$

Der Solomon-Viergruppenplan dient dazu, die mögliche Wirkung von Pretesteffekten zu überprüfen.

## *Kombination der Untersuchungsvarianten (Maaß 2014)*

	Experimentell	Quasiexperimentell
<b>Feld</b>	Interne Validität + Externe Validität +	Interne Validität - Externe Validität +
<b>Labor</b>	Interne Validität + Externe Validität -	Interne Validität - Externe Validität -

Welches Studiendesign wählen Sie zur empirische Evaluation folgender Forschungssoftware?

1. Sie haben eine Weboberfläche entwickelt, die Ressourcen für die Wissenschaftskommunikation bündelt (Zielgruppe: Wissenschaftler:innen)
2. Sie haben eine VR-Plattform entwickelt, mit der Azubis Autos lackieren üben können.
3. Sie haben eine App entwickelt, damit mehr FLINTA Personen Informatik studieren (<https://techstudyfinder.gi.de/>)
4. Sie haben einen Chatbot entwickelt, der auf Social Media versucht zu deeskalieren und produktive Gespräche zu fördern.

- Benoit, Kenneth. 2020. "Text as Data: An Overview." *Annual Review of Political Science* 23: 463–81. <https://doi.org/10.1146/annurev-polisci-051017-020617>.
- Bortz, Jürgen, and Nicola Döring. 2006. *Forschungsmethoden Und Evaluation Für Human- Und Sozialwissenschaftler*. 4th ed. Springer.
- Bubenhofer, Noah, and Joachim Scharloth. 2015. "Korpuslinguistische Diskursanalyse: Der Nutzen Empirisch-Quantitativer Verfahren." *Zeitschrift Für Diskursforschung* 3 (1): 1–24.
- Grimmer, Justin, Margaret E. Roberts, and Brandon M. Stewart. 2022. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press.

Maaz, Kai. 2014. *Quantitative Methoden der Sozialforschung*. University of Potsdam; Lecture slides, University of Potsdam.

Zapf, Antje. 2013. *Empirische Methoden der Sozialforschung*. University of Potsdam; Lecture slides, University of Potsdam.