

**Prof. Dr. Christoph Kulgemeyer**

AG Didaktik der Physik

Fakultät für Naturwissenschaften



# Designs physikdidaktischer Studien



# Wie plant man eine Studie?

---

- Dazu muss man zunächst wissen, was man untersuchen will. Es gibt drei grundlegende Arten:
  - **Explorationsstudien:** Das Ziel hier ist es, neue Theorien zu finden und Hypothesen zu entwickeln. Dazu braucht es ein sogenanntes „theoretisches Sample“.
  - **Survey:** Beschreibt, wie eine bestimmte Gruppe unter einer bestimmten Fragestellung aufgebaut ist - z.B. PISA. Kann auch mit einer Norm vergleichen. Dazu braucht es eine repräsentative Stichprobe.
  - **Evaluationsstudie:** Untersuchung, ob die erwarteten Wirkungen getroffener Maßnahmen auch tatsächlich eingetreten sind: Führt die Einführung der Ganztagschule zu einer höheren Berufszufriedenheit der Lehrpersonen?
  - **Interventionsstudie:** Testet Hypothesen und zugrundeliegende Theorien.

	<b>Survey</b>	<b>Evaluation</b>	<b>Intervention</b>
<b>Vergleichsgruppe Experimentalgr. Kontrollgruppe</b>	ggf. VG	EG	EG + KG od. 2 x EG
<b>Probandenzahl</b>	groß (System Monitoring)	mittel	mittel / klein
<b>Repräsentativität der Stichprobe</b>	sehr wichtig	wichtig	EG und KG müssen vergleichbar sein
<b>Messzeitpunkte</b>	punktuell od. Zeitreihe	<i>formativ:</i> prozessbegleitend <i>summativ:</i> post	prä (zwischen) post follow up

# Drei Arten von Interventionsstudien

- **Nicht-experimentelle Studie:** Beruht auf nicht-randomisierten Stichproben. Problematisch, wenn man kausale Zusammenhänge testen will.
- **Quasi-experimentelle Studie:** Nicht-randomisierte Stichprobe, die aber auf Bedingungen verteilt wird.
- **Experimentelle Studie:** Randomisierte Stichprobe, die auf verschiedene Bedingungen verteilt wird.



# Interne und externe Validität

---

- **Interne Validität:** Eine Studie ist intern valide, wenn man mit hoher Wahrscheinlichkeit annehmen kann, dass der untersuchte Zusammenhang auch gefunden werden kann. Das ist die ganz wesentlich die Validität, wie wir sie bislang kennengelernt haben!
- **Externe Validität:** Eine Studie ist extern valide, wenn die Ergebnisse an verschiedenen Orten und zu verschiedenen Zeiten sowie unabhängig vom Messinstrument gefunden werden können.
  - **Beispiel 1:** Experimentelle Laborstudie in der Universität, um die Wirkung eines Unterrichtsansatzes zu erproben.
  - **Beispiel 2:** Unterrichtsansatz wird an der Schule untersucht, aber in bestehenden Klassenverbänden.

# Stichproben zusammensetzen

---

- **Gelegenheitsstichprobe:** Man nimmt jede Person auf, die teilnehmen möchte

Keine globale Repräsentativität, keine  
variablenspezifische Repräsentativität

- **Quotenstichprobe:** Eine Quote interessanter Variablen bestimmt die Zusammensetzung der Stichprobe

Nur variablenspezifische  
Repräsentativität

- **Randomisierte Stichprobe:** Man weiß, welche Effektstärke man messen will, rechnet aus, wie groß die notwendige Stichprobe ist und nimmt zufällig von allen möglichen Personen diese Anzahl auf.

Globale Repräsentativität und variablenspezifische  
Repräsentativität

# Stichproben zusammensetzen

---

- Häufig in der Physikdidaktik:
  - **Cluster samples (Personen „geclustert“ nach Klassen):**  
Die geclusterte Variable kann zufällig gezogen werden - aber wenn man alle Klassen nimmt, die wollen, ist es nur eine Gelegenheitsstichprobe.

---

**Nach diesen Vorüberlegungen:  
Wie macht man eigentlich schlechte  
Studien?**



# Strategien für schlechte Forschung

**Fall 1:** Sie möchten untersuchen, wie sich eine neue Unterrichtsform auf die Entwicklung des Fachwissens in der Optik auswirkt. Dazu evaluieren Sie ihre eigene Klasse (6 Unterrichtsstunden, Prä und Post) mit kurzen Fragebögen bezüglich des physikalischen Selbstkonzepts. Sie vergleichen diese Ergebnisse mit denen der Klasse Ihres Kollegen, die nach herkömmlichen Verfahren unterrichtet wurde.

- Zwei Hauptprobleme: Design und Operationalisierung
  - Es ist kein experimentelles Design. Man kann nichts über Kausalitäten sagen.
  - Die Testinstrumente müssen das Konstrukt wiedergeben. Hier misst man aber die Selbsteinschätzung der Leistung statt der Leistung.

# Strategien für schlechte Forschung

**Fall 2:** Nun wissen Sie es besser. Sie haben zwei Gruppen mit 200 Schülerinnen und Schülern aus jeweils 10 Klassen gefunden. Sie beschreiben die Intervention detailliert. Nur das Unterrichtsverfahren ändert sich zwischen den Klassen.

- Zwei Probleme: Design und Operationalisierung.
  - Repräsentiert die Stichprobe die Menschen, über die Sie eine Aussage treffen wollen? Ist sie groß genug - oder zu groß?
  - Clustered Sample: Die Klassen variieren systematisch. Das führt zu einer höheren Wahrscheinlichkeit, dass ein beobachteter Effekt falsch ist.

# Strategien für schlechte Forschung

**Fall 3:** Sie wollen wissen, ob direkte Instruktion in Physik einen positiven Einfluss auf die Entwicklung von Fachwissen hat. Dazu re-analysieren Sie Daten aus PISA. Sie vergleichen Länder mit einem hohen Anteil an direkter Instruktion mit solchen, die einen geringen Anteil haben und finden:  $T(2.56)=6.5$  ( $p=.02$ ): ein signifikanter Effekt.

- Zwei Hauptprobleme: Design und Operationalisierung
  - Diese Studie kann keine Entwicklungen abbilden - es ist eine Querschnittstudie, ein Survey.
  - Clusterprobleme?
  - Wie groß ist dieser Effekt eigentlich?