

Empirische Forschungsmethoden I
LMU Wintersemester 2013/14
Dozenten: Karsten Stegmann & Christof Wecker

Janosch Maier

27. Januar 2014

Inhaltsverzeichnis

1 Die Aufgabe pädagogischer Forschung	4
1.1 Wissensbedarf für rationale pädagogische Entscheidungen	4
1.2 Arten pädagogischer Aussagen	4
1.3 Entdeckung und Rechtfertigung pädagogischer Aussagen	4
2 Einführung in die Untersuchungstechniken I: Validität	5
2.1 Begriffsklärung	5
2.2 Bedingungen der Validität von Untersuchungen	5
3 Einführung in die Untersuchungstechniken II: Designs	6
3.1 Untersuchungspläne	6
3.1.1 Within-Subjects-Design	6
3.1.2 Between-Subjects-Design	6
3.2 Darstellung von Untersuchungsdesigns (Campbell & Stanley) . .	7
3.3 Behandlung von Störvariablen	8
4 Einführung in Untersuchungstechniken III: Messung	9
4.1 Dispositionsbegriffe	9
4.2 Operationalisierung	9
4.3 Gütekriterien	9
5 Grundlagen der Hypotheseprüfung	10
5.1 Theorien & Hypothesen	10
5.2 Induktionsproblem & Falsifikation	10
5.3 Signifikanztest	10
5.3.1 Fehlerarten	11
5.3.2 Konventionelle Signifikanzniveaus	11
5.3.3 Ein- & Zweiseitige Tests	11
6 Korrelation	12
6.1 Begriffstypen & Skalenniveaus	12
6.1.1 Klassifikatorische Begriffe	12
6.1.2 Komparative Begriffe	12
6.1.3 Metrische Begriffe	12
6.1.4 Skalensysteme	13
6.2 Systematik inferenzstatistischer Testverfahren	13
6.3 Zusammenhänge zwischen zwei intervallskalierten Variablen: Korrelation	13
6.3.1 Lineare Zusammenhänge	13
6.3.2 Nicht-Lineare Zusammenhänge	13
7 t-Test	14
7.1 t-Test für unabhängige Stichproben	14
7.1.1 Problemtyp	14
7.1.2 Inferenzstatistischer Test	14
7.1.3 Effektstärke	14
7.1.4 Darstellung der Ergebnisse	14
7.2 t-Test für abhängige Stichproben	15
7.2.1 Inferenzstatistischer Test	15

7.2.2	Effektstärke	15
7.2.3	Darstellung der Ergebnisse	15
8	Zweifaktorielle Varianzanalyse	16
8.1	Problemmerkmale	16
8.2	Inferenzstatistischer Test	16
8.3	Interaktionseffekte	17
8.4	Effektstärke	17
9	Gründe für Literaturrecherche	18
10	Fragestellungen & Hypothesen	19
10.1	Fragestellung	19
10.2	Hypothesendreieck	19
10.3	Stichprobenbeschreibung	19
10.4	Design	19
11	Instrumente	21
11.1	Fragebogendaten	21
11.2	Analyse verbaler Daten	21
11.2.1	Codierschema	21
11.2.2	Objektivität	21
11.2.3	Aggregation	21

1 Die Aufgabe pädagogischer Forschung

Beispiel: Einführungsphase im Kindergarten. Kind hängt an Rockzipfel der Mutter und will nicht in Spielecke. Verschiedene Meinungen zwischen Eltern & Erziehern. Eine Realität: Zu vielen Problemen gibt es keine empirischen Daten.

1.1 Wissensbedarf für rationale pädagogische Entscheidungen

Beispiel: Weiterbildung, Kommunikationstraining. Wie wird gelerntes später umgesetzt? Transfer sicherstellen!

- Nicht theoretisches Wissen, sondern Kommunikationstraining vermitteln
- Erfolgreiche Anwendung wird als Transfer bezeichnet
- Goal-based scenarios fördern Transfer und intrinsische Motivation
- Lernende lehnen unbekannte Lernformen ab

Grundsätzlich:

- Welche Ziele? Was vermeiden?
- Ausgangslage?
- Zur Verfügung stehende Maßnahmen?
- Auswirkungen der Maßnahmen auf weitere Bewertungsrelationen?

1.2 Arten pädagogischer Aussagen

- Normative Aussagen (z.B. Ziele)
- Analytische Aussagen (z.B. Definitionen oder direkte Ableitungen)
- Gesetzmäßigkeiten (z.B. Zusammenhänge/Auswirkungen/Allgemeingültig – nur propabilistische Aussagen) – empirisch
- Zustandsbeschreibungen (z.B. Ausgangslage) – empirisch

Empirische Erkenntnisse sind solche, die nur durch Erfahrung (\neq Berufserfahrung sondern, Wahrnehmung/Beobachtungen) möglich sind. (Kant)

1.3 Entdeckung und Rechtfertigung pädagogischer Aussagen

- Idee haben (Entdeckung machen)
- Nachweisen (Rechtfertigen)

Pädagogische Methodenlehre ist eine Argumentationslehre für Aussagen über pädagogisch relevante Sachverhalte.

⇒ Rechtfertigung der Entscheidungen (des Eingangsbeispiels) möglich.

2 Einführung in die Untersuchungstechniken I: Validität

2.1 Begriffsklärung

- Begründe wissenschaftliche Aussagen durch empirische Reihen von Einzelfällen
 - Argument (Beschreibung empirischer Untersuchung von Einzelfällen)
 - ⇒ (Validität? – These als Schluss aus Argumenten zutreffend?)
 - These (Wissenschaftliche Aussage)
- Empirische Untersuchung ist **valid** gdw. Darstellung gültiges Argument zur untermauernden Hypothese ist
- Unabhängige Variable (Ursache, wird in Experiment manipuliert)
- Abhängige Variable (Wirkung, wird in Experiment gemessen)

2.2 Bedingungen der Validität von Untersuchungen

- Untersuchungsvalidität
 - Eliminationsbedingung / **Interne Validität**: Elimination anderer Ursachen bei kausalen Hypothesen. Ausschluss von Kofundierung (Variable, die mit unabh. Variable gemeinsam variiert) ⇒ Wird durch entsprechendes Versuchsdesign ausgeschlossen
 - Repräsentativitätsbedingung / **Externe Validität**: Repräsentativität der Studie (Zufallsstichproben, Stichproben aus Population, Stichproben aus Geltungsbereich der Hypothes) ⇒ Wird durch entsprechende Stichprobe ausgeschlossen
 - Statistische Bedingungen / Statistische Validität: Korrekte statistische Berechnung
- Variablenvalidität
 - Bedingung der Validität der Zuweisung von Werten auf Variablen / Validität der **Operationalisierung**: Anwendung der Testtheorie, Aufteilung in (un-)abhängige Variablen

3 Einführung in die Untersuchungstechniken II: Designs

3.1 Untersuchungspläne

- Design: Plan, an welchen Versuchseinheiten welche Variablen manipuliert bzw. gemessen werden
- Treatment: Ausprägung einer unabhängigen Variable

3.1.1 Within-Subjects-Design

- Jede VP wird unter verschiedenen experimentellen Bedingungen untersucht (UV an-/abwesend)
- Vorteile:
 - Weniger VP benötigt
 - Kontrolle personengebundener Störvariablen
 - Direkteste Untersuchung, Welche UV prägt AV am stärksten aus
- Probleme:
 - Lange Untersuchungssitzungen
 - Fehlende Vergleichbarkeit von Messungen
 - Übertragungseffekte (1. Messung kann 2. Messung beeinflussen). z.T. kontrollierbar durch Balancierung (Tauschen der Reihenfolge)

3.1.2 Between-Subjects-Design

- Jede VP wird nur unter einer experimentellen Bedingung untersucht.
- Experiment: Zufällige Aufteilung auf experimentelle Bedingungen
- Quasi-Experiment: Aufteilung von vorgefundenen Gruppen auf experimentelle Bedingungen

1. Pre-Post-Kontrollgruppenplan

- Kontrollgruppe: Pretest, Posttest
- Experimentalgruppe: Pretest, Treatment, Posttest
- Vorteile:
 - Unterschiede im Posttest auf Treatment zurückführbar
 - Vergleichbarkeit zwischen Gruppen überprüfbar
- Nachteile:
 - Unterschied tritt evtl. nur auf, wenn Pretest durchgeführt

2. Kontrollgruppenplan

- Kontrollgruppe: Posttest
- Experimentalgruppe: Treatment, Posttest

- Vorteile:
 - Unterschiede im Posttest auf Treatment zurückföhrbar
 - Auswirkung von Prättest ausgeschlossen
- Nachteile:
 - Vergleichbarkeit von Gruppen nicht überprüfbar

3. Solomon-Vier-Gruppen-Plan

- Kontrollgruppe 1: Pretest, Posttest
- Experimentalgruppe 1: Pretest, Treatment, Posttest
- Kontrollgruppe 2: Posttest
- Experimentalgruppe 2: Treatment, Posttest
- Vorteile:
 - Unterschiede im Posttest auf Treatment zurückföhrbar
 - Auswirkung von Prättest ausgeschlossen
- Nachteile:
 - Schwierig auszuwerten

4. Mehr-Gruppen-Plan

- Kontrollgruppe: Pretest, Posttest
- Experimentalgruppe 1: Pretest, Treatment 1, Posttest
- Experimentalgruppe 2: Pretest, Treatment 2, Posttest
 - Vergleich mehrerer Äusprägungen von UV möglich
 - Mit oder ohne Pretest durchführbar

5. Mehrfaktorielle Pläne

- Gruppe 1: Pretest, UV 1, UV 2 Posttest
- Gruppe 2: Pretest, UV 1, Posttest
- Gruppe 3: Pretest, UV 2, Posttest
- Gruppe 4: Pretest, Posttest
- Vorteile:
 - Gleichzeitige Untersuchung mehrerer UV
 - Untersuchung eines Treatments unter verschiedenen Bedingungen

3.2 Darstellung von Untersuchungsdesigns (Campbell & Stanley)

(Zeitlicher Ablauf)	Pretest	Treatment	Posttest
Gruppe 1	o (Messung vorhanden)	x (Treatment vorhanden)	o
Gruppe 2	(Messung nicht vorhanden)	(Treatment nicht vorhanden)	

3.3 Behandlung von Störvariablen

Variablen, die nicht in Hypothese vorkommen, welche AVs beeinflussen können

- Konstanthalten: Nur Untersuchung, wenn Kontrollvariable (Störvariable, die kontrolliert wird) dieselbe Ausprägung (Einfluss ausgeschlossen, aber keine Verallgemeinerung möglich)
- Systematische Variation: Untersuchung bei unterschiedlicher Ausprägung der Kontrollvariable (Einfluss ausgeschlossen, aber keine Verallgemeinerung möglich)
- Parallellisierung und Matching: Gleicher Mittelwert der Störvariable / Paare von Personen mit gleicher Ausprägung gebildet, dann zufällig auf Gruppen aufgeteilt (Einfluss ausgeschlossen, Vorherige Messung der SV nötig)
- Randomisierung: Zufällige Aufteilung in Gruppen (Keine Messung nötig, alle SV kontrollierbar (Königsweg), aber nicht überprüfbar, wenn keine Messung)
- Statistische Kontrolle: Nachträgliches Herausrechnen der SV (Nachträglich möglich, aber Messung nötig, allerdings nicht unbedingt vor Gruppeneinteilung)
- Pretest ist **keine** Störvariable, da er in allen Versuchsdesigns vorkommt.

4 Einführung in Untersuchungstechniken III: Messung

4.1 Dispositionsbegriffe

- Was soll erfasst werden? Psychologische Eigenschaften wie Wissen, Selbstkonzept, ...
- Standardbeispiel: Wasserlöslich
 - Lässt sich nicht direkt erkennen (“Nicht durch ansehen beantwortbar”)
 - Zur Feststellung muss ein Test durchgeführt werden (“in Wasser legen”)
 - Testergebnis zeigt Disposition (“Kein Pulver übrig \Rightarrow Wasserlöslich”)

4.2 Operationalisierung

- Operation zur Erfassung des Sachverhaltes
- Angabe von Messverhalten, die Sachverhalt anzeigen
- Messung oder Manipulation sind Operationalisierungen

4.3 Gütekriterien

- Objektivität: Unabhängigkeit vom Versuchsleiter. Gegeben, wenn unterschiedliche Personen zum gleichen Ergebnissen kommen.
 - Durchführungsobjektivität: Genaue Beschreibung des Versuchs
 - Auswertungsobjektivität: Regeln / Kodiersystem für Auswerter
 - Interpretationsobjektivität: Regeln zur Interpretation
- Reliabilität
 - Zuverlässigkeit wie genau ein Wert gemessen wird.
- Validität
 - Ausmaß, in dem der Test das misst, was er messen soll.
 - * Inhaltlich: Direkt am Test ist sichtbar, was er misst
 - * Kriteriumsbezogen: Übereinstimmung des Test mit anderen Kriterien (Innere Validität: Vergleich mit Test, der selbe Variable misst; Äußere Validität: Vergleich mit objektivem Maß [Physiologie] oder Expertenrating)
 - * Konstrukt: Testverhalten im Einklang mit Erwartungen über teilnehmende Personen

Objektivität Voraussetzung für Reliabilität, Voraussetzung für Validität

5 Grundlagen der Hypotheseprüfung

Theorie / Hypothese → Empirie → Entscheidungen → Theorie / ...

5.1 Theorien & Hypothesen

- Theorie muss theoretisch widerlegbar sein
 - Sollte logisch widerspruchsfrei
 - Sollte möglichst viele Beobachtungen erklären
 - Sollte konkrete Vorhersagen ableiten lassen
- Hypothese postuliert Zusammenhang zwischen mindestens zwei Begriffen
 - Bezug auf reale, empirische untersuchbare Sachverhalte
 - Allgemeingültig, geht über Einzelfall hinaus
 - Implizite Formalstruktur (wenn, dann)
 - Widersprechende Ereignisse müssen denkbar sein
 - Unabhängige Variablen bestimmen abhängige Variablen hervor

5.2 Induktionsproblem & Falsifikation

- Induktion Schluss von Einzelfällen (Prämisse) auf allgemeine Gesetzmäßigkeit (Konklusion) logisch ungültig. Immer nur begrenzte Versuchspersonenanzahl, aber Aussagen über unbegrenzt große Mengen.
- Kritischer Rationalismus: Popper – Verifizierbarkeit nicht endgültig möglich
- Wenn Experiment Aussage bestätigt, hat sie sich bewährt

5.3 Signifikanztest

- Hypothese verwerfen auf Grundlage der Daten
- H_1 ist Alternativhypothese, die mich interessiert
- H_0 ist Nullhypothese, aus der Aussagen über die Wahrscheinlichkeit von Daten abgeleitet werden
- Verwerfen der Nullhypothese, wenn Daten bei Geltung von H_0 extrem unwahrscheinlich ist (Keine Widerlegung)
- **Hypothesen haben keine Wahrscheinlichkeiten! Kann von der Statistik nicht geleistet werden.**

5.3.1 Fehlerarten

- Tatsächlich zutreffend: H_0
 - Entschieden für H_0 : richtig
 - Entschieden für H_1 : α -Fehler (1. Art) – Nullhypothese wird verworfen, obwohl sie gilt (Muss unbedingt minimiert werden).
- Tatsächlich zutreffend: H_1
 - Entschieden für H_0 : β -Fehler (2. Art) – Nullhypothese wird nicht verworfen, obwohl Alternativhypothese gilt.
 - Entschieden für H_1 : richtig

5.3.2 Konventionelle Signifikanzniveaus

- $< 1\%$: Hochsignifikant (**)
- $< 5\%$: Signifikant (*)
- häufig $< 10\%$: “tendenziell” (marginal effect) – NICHT VERWENDEN

5.3.3 Ein- & Zweiseitige Tests

- Gerichtete Hypothesen:
 - Wenn ... ist ... höher/niedriger, als wenn ...
 - Je mehr, desto mehr/weniger ... (positiver / negativer Zusammenhang)
- Ungerichtete Hypothesen
 - Es besteht ein Unterschied zwischen ... und ... hinsichtlich ...
 - Es besteht ein Zusammenhang zwischen ... und ...
- Einseitiger Test: Testen einer gerichteten Hypothese.
- Zweiseitiger Test: Testen einer ungerichteten Hypothese.

6 Korrelation

6.1 Begriffstypen & Skalenniveaus

- Dimension: Menge zusammengehöriger Begriffe. Objekte einer Dimension haben einen bestimmten Begriffstyp
- Ausprägung: Elemente einer Dimension

6.1.1 Klassifikatorische Begriffe

- Einteilung einer Gegenstandsmenge in Klassen. Bsp: Schultyp (Hauptschule, Realschule, ...)
- Adäquatheitsbedingungen: Vollständige Zerlegung / Erschöpfung (mindestens ein Begriff zutreffend), Ausschließlichkeit (maximal ein Begriff zutreffend) \Rightarrow genau ein Klassifikatorischer Begriff zutreffend für ein Objekt
- Dichotom (Zwei Ausprägungen) / Polytom (Mehrere Ausprägungen)
- Nominalskala: Umbenennung der Werte (z.B. in Nummern), Mögliche Transformationen: (Ein-)Eindeutige, Aussagen: Gleichheit / Verschiedenheit

6.1.2 Komparative Begriffe

- Paarweise Einordnung in Reihenfolge (Mehr; mindestens so ..., wie)
- Ordinal- / Rangskala (Reihenfolge), Mögliche Transformationen: Monoton (Reihenfolge muss erhalten bleiben), Aussagen: Relative Rangordnung

6.1.3 Metrische Begriffe

- Zuordnungen von Zahlen für Objekte zum Rechnen. Bsp: Intelligenz, Klassengröße, Alter
- Intensiv (Operation: Zusammenfügen ergibt Summe – nicht möglich. Bsp: IQ) vs. Extensiv (Operation: Zusammenfügen ergibt Summe – möglich. Bps: Zeit)
- Intervallskala (intensiv), Mögliche Transformationen: Lineare Transformation $x' = ax + b$, Aussagen: Gleichheit von Differenzen
- Rationalskala (extensiv), Mögliche Transformationen: Ähnlichkeitstransformationen $x' = ax$, Aussagen: Gleichheit von Verhältnissen
- Differenzskala, Mögliche Transformationen: $x' = x + c$
- Absolutskala, Mögliche Transformationen: $x' = x$ (Identität). Bsp: Wahrscheinlichkeit

6.1.4 Skalensysteme

- Versuch komparative Begriffe auf Intervallskalen abzubilden
- Größer-Kleiner Beziehung / Reihenfolge? Nein → Nominalskala
- Gleiche Abstände? Nein → Ordinalskala
- Natürlicher, sinnvoller Nullpunkt? Nein → Intervallskala
- Ja → Rationalskala

6.2 Systematik inferenzstatistischer Testverfahren

Sehr wichtig!

UV(s)	# UVs	# Faktorstufen	Stichproben	AV: Intervallskaliert
nominal	1	2	unabhängig	t-Test für unabhängige Stichproben
			abhängig	t-Test für abhängige Stichproben
		>2	unabhängig	einfaktorielle Varianzanalyse
	>1		unabhängig	mehrfaktorielle Varianzanalyse
intervall				Korrelation

6.3 Zusammenhänge zwischen zwei intervallskalierten Variablen: Korrelation

6.3.1 Lineare Zusammenhänge

- Streudiagramm / Scatterplot (Gerade)
- Dreidimensionaler Graph einer bivariaten Dichtefunktion (Gauß-Verteilung)
- Deskriptiver Kennwert: Pearsonser Korrelationskoeffizient / Produkt-Moment-Korrelation r_{xy} – Wertebereich $[-1, 1]$: -1 = perfekter negativer Zusammenhang, 0 = Kein Zusammenhang, 1 = perfekter positiver Zusammenhang
- Effektstärkenmaß: Determinationskoeffizient r^2 – Anteil der Varianz einer Variablen, die durch andere aufgeklärt wird. 0,01 = Kleiner Effekt, 0,09 = mittlerer Effekt, großer Effekt = 0,25
- Inferenzstatistischer Test
 - Verteilung von Korrelationskoeffizienten in Stichproben aus einem Universum ohne Zusammenhang (Nullhypothese) folgen der t-Verteilung
 - Zu festgestelltem Korrelationskoeffizient wird t -Wert bestimmt. Berechnung der Wahrscheinlichkeit, dass unter Geltung der Annahme kein Zusammenhang besteht
 - Wenn Wahrscheinlichkeit niedriger als Signifikanzniveau, wird Nullhypothese verworfen. \Rightarrow Zusammenhang besteht.

6.3.2 Nicht-Lineare Zusammenhänge

- z.B. Parabelförmige Kurve
- Korrelationskoeffizient = 0

7 *t*-Test

7.1 *t*-Test für unabhängige Stichproben

7.1.1 Problemtyp

Siehe Tabelle.

- UV: Nominalskaliert
- AV: Intervallskaliert
- #UV: 1
- #Faktorstufen der UV: 2
- Unabhängige Stichproben

7.1.2 Inferenzstatistischer Test

- Nullhypothese: “Mittelwerte unterscheiden sich nicht”
- Grundprinzip:
 - *t*-Wert berechnen: Wahrscheinlichkeit, dass Stichproben sich unterscheiden, wenn Nullhypothese gilt (Universen sind gleich)
 - Wenn Wahrscheinlichkeit < 1% oder < 5%, Nullhypothese verwerfen

7.1.3 Effektstärke

- Effektstärke *d*: Kleiner Effekt: 0,2 / Mittlerer Effekt: 0,5 / Großer Effekt: 0,8
- $d \sim \frac{\text{Mittelwert}}{\text{Streuung}}$

7.1.4 Darstellung der Ergebnisse

- Grundsätzlich:
 - Deskriptiv: Ergebnis in Stichprobe – Darstellung als Text (unbedingt!) + Tabelle oder Grafik
 - Inferenzstatistisch: Verallgemeinerung möglich?
 - * Prüfstatistik: *t*-Wert. $t_{\text{Freiheitsgrade}} = t\text{-Wert}$
 - * Signifikanzwert: $p < \text{Signifikanzwert}$, α -Fehler / n.s. (nicht Signifikant)
 - * Effektstärke: $d = \text{Effektstärke}$

7.2 *t*-Test für abhängige Stichproben

Siehe Tabelle.

- UV: Nominalskaliert
- AV: Intervallskaliert
- #UV: 1 (Kein Treatment – Pretest / Treatment – Posttest)
- #Faktorstufen der UV: 2
- **Abhängige** Stichproben

7.2.1 Inferenzstatistischer Test

analog zum t-Test für unabhängige Variablen

- Nullhypothese: “Mittelwerte unterscheiden sich nicht”
- Grundprinzip:
 - *t*-Wert berechnen: Wahrscheinlichkeit, dass Stichproben sich unterscheiden, wenn Nullhypothese gilt (Universen sind gleich)
 - Wenn Wahrscheinlichkeit < 1% oder < 5%, Nullhypothese verwerfen

7.2.2 Effektstärke

- Effektstärke *d*: Kleiner Effekt: 0,2 / Mittlerer Effekt: 0,5 / Großer Effekt: 0,8
- $d \sim \frac{\text{Mittelwert}}{\text{Streuung}}$

7.2.3 Darstellung der Ergebnisse

- Grundsätzlich:
 - Deskriptiv: Ergebnis in Stichprobe – Darstellung als Text (unbedingt!) + Tabelle oder Grafik
 - Inferenzstatistisch: Verallgemeinerung möglich?
 - * Prüfstatistik: *t*-Wert. $t_{\text{Freiheitsgrade}} = t\text{-Wert}$
 - * Signifikanzwert: $p < \text{Signifikanzwert}$, α -Fehler / n.s. (nicht Signifikant)
 - * Effektstärke: $d = \text{Effektstärke}$

8 Zweifaktorielle Varianzanalyse

8.1 Problemmerkmale

- Mehrere UVs (Nominalskaliert), beliebige Faktorstufen, unabhängige Stichprobe (Faktorstufen bei unterschiedlichen Personengruppen)
- AVs (Intervallskaliert)
- Typisches Problem: Aptitude-treatment interactions (ATI-Effekt). Unterschiedliche Wirksamkeit von Treatments bei Personen mit unterschiedlichen Merkmalen (aptitudes)

- Darstellung als 2x2-Matrix:

		Belohnung	
		mit	ohne
Geschlecht	♂		
	♀		

8.2 Inferenzstatistischer Test

- ANOVA: Varianzanalyse (Analysis of Variance)
- Faktor (factor): unabhängige Variable (Independent Variable)
- Faktorstufe: Ausprägung der unabhängigen Variable
- Faktorstufenkombination (Matrix-Zelle): Kombination von UV-Ausprägungen
- Haupteffekt (main effect): Unterschied der Mittelwerte der AV bei verschiedenen Ausprägungen einer UV. Bsp für H_0 : Mittelwerte der Faktorstufen unterscheiden sich nicht.
- Interaktionseffekt (interaction effect): Unterschiedsmuster der Mittelwerte der AVs bei unterschiedlicher Faktorstufenkombination (Höhere Abweichung der MWs einer AV bei einer bestimmten Ausprägung der anderen UV). Interaktionseffekt möglich, auch wenn keine Haupteffekte vorliegen. Bsp für H_0 : Mittelwerte der Faktorstufenkombinationen setzen sich additiv aus den Haupteffekten zusammen
- Moderatorvariable (moderator variable): Variable, von der Effekt einer anderen Variable abhängt

Einschub: Klausurvorbereitung – Mediator / Moderator

- Mediator: Vermittelnde Variable – Steht zwischen UV und AV (Wird von UV beeinflusst, beeinflusst AV) – Ist UV (in einem Teil) und AV (in anderem Teil)
- Moderator: Verändert den Zusammenhang zwischen zwei Variablen – vgl. Interaktionseffekt bei mehrdimensionalen Designs – Ist UV (Ausnahme statistisches Rausrechnen – *für uns unwichtig*)

8.3 Interaktionseffekte

vgl. Skript Seite 33.

- Kein Interaktionseffekt (Additionseffekt. Gleiche Steigung).
- Ordinale Interaktion (Unterschiedliche Steigung). Alle Haupteffekte gelten.
- Disordinale Interaktion (Negative Steigung einer Geraden auf beiden Bildern). Haupteffekte gelten nicht mehr.
- Hybride Interaktion = ein Ordinale, ein Disordinaler Haupteffekt. (Negative Steigung einer Geraden einem Bild). Ein Haupteffekt gilt

8.4 Effektstärke

- Effektstärkenmaß: η^2 ("Eta-Quadrat"). Anteil der Variabilität, der auf Faktorstufen zurückgeht (Erklärte Varianz)
- Konventionelle Grenzwerte
 - Kleiner Effekt: $\geq 0,01$
 - Mittlerer Effekt: $\geq 0,06$
 - Großer Effekt $\geq 0,14$

9 Gründe für Literaturrecherche

vgl. Folie 36. "Literaturstudium als Grundlage empirischer Studien"

- Internet < Lehrbücher < Fachstudien \Rightarrow Am Ende alle relevante Literatur kennen
- Handbücher statt Lehrbücher verwenden
- Wichtig: Vermeiden, relevante Studien zu übersehen
- Fachbegriffe, Theoretische Ansätze, Empirische Befundlage, Methodische Ansätze \Rightarrow Neue Informationssuche
- Konsistentes Modell. Neue Texte lassen sich konfliktfrei integrieren \Rightarrow Planung empirischer Studie konkretisieren, aber Weiterlesen und Planung anpassen, Konkrete Planung führt zu neuen Fragen.
- Kontrolle: Lösen intendierte Erkenntnisse ursprüngliches Problem?

10 Fragestellungen & Hypothesen

10.1 Fragestellung

- In wie weit haben UV1 (ohne, mit) und UV2 (ohne, mit) und deren Interaktion
- unabhängig von einer KV
- einen Effekt auf AV
- im Kontext

10.2 Hypothesendreieck

- RQ = Research Question (Fragestellung)
- In wie weit hat die UV (ohne, mit) einen Effekt auf ... (RQ1, RQ3)
- In wie weit besteht ein Zusammenhang zwischen AV1 (MV) und AV2 (AV)
- Eine Fragestellung enthält meist mehrere Hypothesen

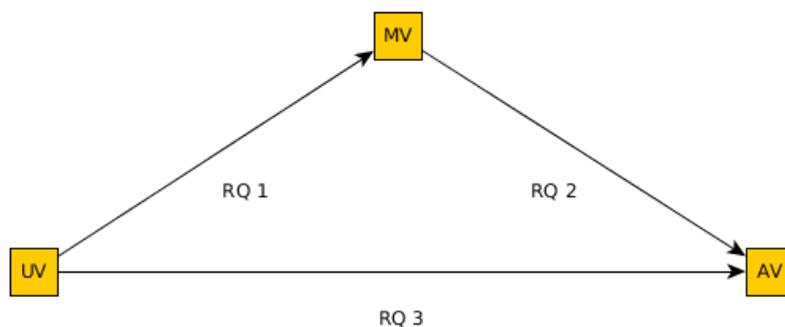


Abbildung 1: Hypothesendreieck

10.3 Stichprobenbeschreibung

- Alter, Geschlecht angeben
- z.B. Mittelwert, Standardabweichung,

10.4 Design

- 4-Felder-Tafel
- Bericht, wie viele Probanden welche Merkmale der UVs zugewiesen bekamen
- Feld- / Laborexperiment, ...

- Kofundierung beachten (Quasi-Experiment) – z.B. Teilnahme am Tutorium kofundiert mit Motivation
- Kontrolltechniken: Pre-Post-Test, Balancierung, Systematisches Variieren, Variablen Messen, Randomisierung
- Wie wird umgesetzt? Operationalisierung der unabhängigen Variablen

11 Instrumente

11.1 Fragebogendaten

- Teilweise Problematisch, da Selbstauskünfte (Soziale Erwünschtheit)
- Einfach auszuwerten
- Kronbach's Alpha: Interne Konsistenz – Reliabilität
 - $\alpha \geq 0.9$ – Exzellent
 - $0.8 \leq \alpha < 0.9$ – Gut
 - $0.7 \leq \alpha < 0.8$ – Akzeptabel
 - $0.6 \leq \alpha < 0.7$ – Ausreichend
 - $0.5 \leq \alpha < 0.6$ – Gering
 - $\alpha < 0.5$ – Nicht akzeptabel
- Anzahl der Items
- Beispiel eines Items
- Reliabilität
- (Mittelwert und Standardabweichung über alle Versuchspersonen)

11.2 Analyse verbaler Daten

11.2.1 Codierschema

- Prozessanalysen vs. Ergebnisanalysen
- Kodiersystem (z.B. Propositionstyp) verwenden
- Segmentierung, wenn Anzahl benötigt
- Keine Segmentierung, wenn nur vorhanden / nicht vorhanden

11.2.2 Objektivität

- Auswertungsobjektivität: Mehrere Leute auswerten lassen (Alle kodierten Codes müssen vorkommen \Rightarrow 10% der Daten kodieren), dann vergleichen (Paarweise vergleichen)
- Problematische Kategorien “kollabieren”
- Cohens Kapa:
 - 0.4: Notfalls Akzeptabel
 - 0.6: Ausreichend
 - 0.8: Gut

11.2.3 Aggregation

- Absolute Anzahl: Vorkommen der Ausprägungen einer Kategorie
- Relativer Anteil: Vorkommen der Ausprägungen einer Kategorie

11.3 Zusammenfassung

- Fragebögen / Tests: Ergebnisse Codieren und Aggregieren (Tabelle)
- Beobachtungsdaten: Transkription, Segmentierung, Codierung, Aggregation (Objektivität mit Cohens Kappa prüfen)