

Grundlagen Datenerhebung  
LMU Wintersemester 2011/12  
Dozent: Thomas Stoffer

Janosch Maier

26. Januar 2012

Basierend auf dem Skript zur Vorlesung Grundlagen Datenerhebung.  
[www.psy.lmu.de/exp/teaching/courses/index.html](http://www.psy.lmu.de/exp/teaching/courses/index.html). Bilder entsammen von dort.

## Inhaltsverzeichnis

<b>1</b>	<b>Wissenschaftstheoretische Grundlagen</b>	<b>3</b>
1.1	Psychologie als Wissenschaft . . . . .	3
1.2	Erkenntnisgewinnung in der Psychologie . . . . .	3
1.2.1	Prinzipien der Erkenntnisgewinnung . . . . .	3
1.2.2	Wissenschaftliche Hypothesenprüfung . . . . .	3
<b>2</b>	<b>Methodologische Ansätze</b>	<b>4</b>
2.1	Hypothese generierung und Operationalisierung . . . . .	4
2.1.1	Gütekriterien für Hypothesen . . . . .	4
2.1.2	Operationalisierung von Variablen . . . . .	4
2.2	Korrelatives / Experimentelles Untersuchungsdesign . . . . .	4
2.2.1	Kovariation von Variablen . . . . .	4
2.2.2	Korrelatives Untersuchungsdesign . . . . .	5
2.2.3	Experimentelles Untersuchungsdesign . . . . .	5
2.3	Quantitative vs. Qualitative Operationalisierung . . . . .	5
2.3.1	Methodenwahl und Forschungsfortschritt . . . . .	5
2.3.2	Qualitative vs. Quantitative Methoden . . . . .	5
2.3.3	Menschenbildimplikation . . . . .	5
<b>3</b>	<b>Kriterien Psychologischer Forschung</b>	<b>6</b>
<b>4</b>	<b>Systematische Verhaltensbeobachtung</b>	<b>6</b>
4.1	Wissenschaftliche Beobachtung . . . . .	6
4.2	Strukturierte / Standardisierte Beobachtung . . . . .	7
4.2.1	Beobachtungsfehler . . . . .	7
4.2.2	Kategoriesysteme . . . . .	8
4.2.3	Reliabilität der Beobachtung: Beobachtungsübereinstimmung . . . . .	8

<b>5</b>	<b>Interview und Fragebogen</b>	<b>9</b>
5.1	Beobachtung, Fragebogen oder Interview? . . . . .	9
5.2	Arten und Ziele von Fragebogen . . . . .	9
5.2.1	Grad der Standardisierung . . . . .	9
5.2.2	Hauptanwendungsgebiete für Fragebogen . . . . .	10
5.3	Schritte der Fragebogenkonstruktion . . . . .	10
5.3.1	Inhaltliche Festlegung . . . . .	10
5.3.2	Festlegung des Fragentyps . . . . .	10
5.3.3	Formulierung von Fragen und Antwortkategorien . . . . .	12
5.3.4	Reihenfolge der Fragen . . . . .	12
5.3.5	Erprobung und Überarbeitung des Fragebogenentwurfs . . . . .	12
<b>6</b>	<b>Experimentelle Versuchsmethodik</b>	<b>13</b>
6.1	Das Experiment . . . . .	13
6.2	Variablentypen . . . . .	14
6.3	Prinzipien der Operationalisierung . . . . .	14
6.3.1	Operationalisierung UVs . . . . .	14
6.3.2	Operationalisierung AVs . . . . .	15
6.4	Probleme bei Messung von AVs . . . . .	15
6.4.1	Neutrale Messung bei Selbst- und Fremdbeobachtung . . . . .	15
6.4.2	Reliabilität von Messungen / Messgenauigkeit . . . . .	15
6.5	Probleme der Validität eines Experiments . . . . .	15
6.5.1	Experimentelle Kontrolle . . . . .	15
6.5.2	Sicherstellung interner Validität . . . . .	16
6.5.3	Sicherstellung externer Validität . . . . .	16
6.5.4	Labor- vs. Feldexperiment . . . . .	17
<b>7</b>	<b>Versuchspläne</b>	<b>18</b>
7.1	Experimentelle Versuchsplanung . . . . .	18
7.1.1	Ziele der Versuchsplanung . . . . .	18
7.2	Quasiexperimentelle Versuchspläne und ihre Erweiterung zu experimentellen Versuchspläne . . . . .	18
7.2.1	Einmalige Nacher-Messung an einer Stichprobe . . . . .	18
7.2.2	Vorher-Nacher-Messung an einer Stichprobe . . . . .	18
7.2.3	Statistischer Vergleich vorgegebener Gruppen . . . . .	18
7.3	Experimentelle Designs . . . . .	19
7.3.1	Randomisierte Versuchspläne . . . . .	19

# 1 Wissenschaftstheoretische Grundlagen

## 1.1 Psychologie als Wissenschaft

Wissenschaft:

- Gegenstand
- Methoden
  - Systematik
  - Rationalität
  - Objektivität

## 1.2 Erkenntnisgewinnung in der Psychologie

### 1.2.1 Prinzipien der Erkenntnisgewinnung

Wissenschaftliche Methoden:

- Intuitives Wissen kritisch hinterfragen
- Wissenschaftliches Wissen überprüfen
- Übereinstimmung, Widerspruchsfreiheit von Erkenntnissen prüfen
- Logische Fehler (bei der Beobachtung) in Theorien vermeiden

Erkenntnisgewinnung durch:

- Prinzip der Vernunft (Formulierung)
- Prinzip der Erfahrung (Überprüfung)

Früher: Hermeneutik

- Deutung von einzelnen Verhaltenserscheinungen, Hineinversetzen
- Deutung meist unzutreffend!

Skeptizismus notwendig!

- Intuition / Autorität fehleranfällig
- Gefahr von Schädigung / Wirkungslosigkeit

### 1.2.2 Wissenschaftliche Hypothesenprüfung

**Theorien an der Realität prüfen** ↔ Vergleich von Vorhersagen mit Beobachtungen

Induktive Schlüsse schwierig (Allaussagen!), da leicht zu falsifizieren.

#### **Falsifikationsprinzip**

- Deduktiv, der Theorie widersprechende Nullhypothese ableiten
- Nullhypothese durch Beobachtungen falsifizieren

- Erfolgreiche falsifikation impliziert richtigkeit der Theorie, sonst Theorie belastet

Exhaustion einer belasteten Theorie: Erklärung der Daten, durch Modifikation der Theorie / Aufzeigen von Störvariablen

Überprüfung von Theorien (Rechtfertigung / Entdeckung)

- Theorie / Entdeckung
- Hypothesen / Beschreibung
- Operationalisierung
- Verhaltensvorhersagen
- Datenerhebung
- Statistische Analyse
- Schlussfolgerung

## 2 Methodologische Ansätze

### 2.1 Hypothese generierung und Operationalisierung

#### 2.1.1 Gütekriterien für Hypothesen

- Testbarkeit
- Falsifizierbarkeit
- Präzisierung
- Theorie relevanz

#### 2.1.2 Operationalisierung von Variablen

- Definition von Begriffen
- Übersetzen in beobachtbare Erlebens / Verhaltensmerkmale

Verankerung der Hypothesenkonzepte in der Realität

Resultat:

- Unabhängige Variablen (UV)
- Abhängige Variablen (AV)

### 2.2 Korrelatives / Experimentelles Untersuchungsdesign

#### 2.2.1 Kovariation von Variablen

Kovariation: Zusammenhangsstärke zwischen (mind. 2) Variablen. Regressionsgerade ist lineare Funktion.

Korrelation: Mittlere Abweichung von der Regressionsgeraden.

Hypothesen behaupten Kovariation von Variablen. Möglicherweise Ursache (UV), Wirkung (AV).

## **2.2.2 Korrelatives Untersuchungsdesign**

Korrelationsstudien erfassen nur Kovariation (keine Kausalität)

Korrelationsstudien, wenn Manipulation von UV nicht möglich. Versuchsbedingung durch Selektion.

## **2.2.3 Experimentelles Untersuchungsdesign**

- Theoretisch begründete, kausale Hypothese
- Herstellung & Manipulation des Verhaltens
- Experimentelle Kontrolle (Siehe Kapitel 6.5.1)

## **2.3 Quantitative vs. Qualitative Operationalisierung**

### **2.3.1 Methodenwahl und Forschungsfortschritt**

Hypothese generierung: Qualitative & Quantitative Methoden

Hypothesenprüfung: Nur Quantitative Methoden

### **2.3.2 Qualitative vs. Quantitative Methoden**

Quantitativ:

- Beobachtungen bilden Verhalten auf Zahlen ab
- Quantifizierung
- Vergleich von Verhaltensphänomenen
- Zusammenfassung von Ergebnissen
- Unterscheidung von Zufallsbeobachtungen von systematischen Tendenzen
- Erfassen von Messfehler (Genauigkeit) & Reliabilität (Zuverlässigkeit)
- Präzise Hypotheseüberprüfung

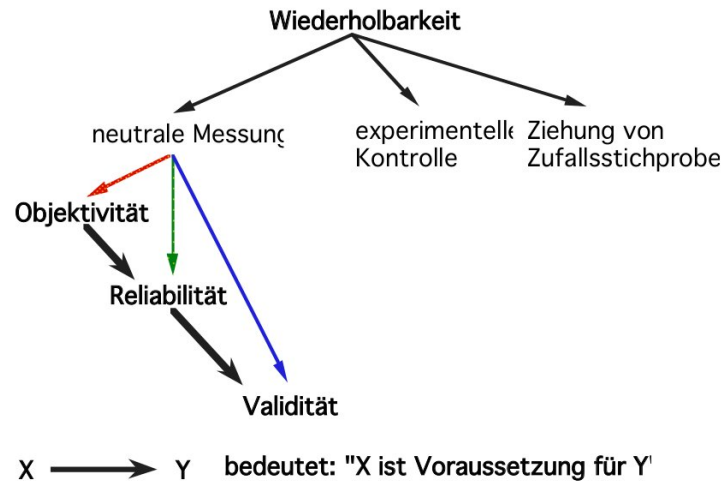
Qualitativ:

- Transformation in Zahlen unangemessen
- Methoden mit Selbstauskünften (Introspektion, teilnehmende Beobachtung)

### **2.3.3 Menschenbildimplikation**

Quantitative Methoden reduzieren Menschen auf Zahlen  $\leftrightarrow$  Zahlen sind nur Hilfsmittel (Nur zur Überprüfung)

### 3 Kriterien Psychologischer Forschung



- Wiederholbarkeit: Experiment unter gleichen Experimentalbedingungen beliebig oft Wiederholbar. Absicherung / Vergleich. Nicht (optimal) gegeben bei: Introspektion, teilnehmender Beobachtung
- Neutrale Messung: Keine Beeinflussung des Ergebnisses durch Messung (Bias, Reaktivität)
  - Objektivität: Intersubjektive Überprüfbarkeit (Verschiedene Personen kommen unabhängig von einander auf selbes Ergebnis)
    - \* Darbietungsobjektivität (genau gleiche Darbietung)
    - \* Auswertungsobjektivität (genau gleiche Auswertung)
    - \* Interaktionsobjektivität (Interpretation nur auf Grund von Statistik)
  - Reliabilität: Genauigkeit der Messung (Gleiche Messwerte bei Wiederholung der Messung)
  - Validität: Messung der Richtigen Variablen
- Experimentelle Kontrolle: Kontrolle von Störvariablen
- Ziehung von Zufallsstichproben: Nur zufällige Unterscheidung von Grundgesamtheit (Gruppe, auf die generalisiert werden soll)

### 4 Systematische Verhaltensbeobachtung

#### 4.1 Wissenschaftliche Beobachtung

Voraussetzungen für Wissenschaftlichkeit:

- Spontanes Verhalten, das ausreichend häufig auftritt
- Beobachtung zur prüfung einer Hypothese
- Wissenschaftliche Kriterien

- Minimierung von Reaktivität (Messfehler verletzt Objektivität)
- Repizierbarkeit (Reliabilität)
- Richtiges Verhalten erfassen (Validität)

Beobachtung ist Methode der Datenerhebung, nicht des Forschungsdesigns. Möglich bei korrelativer & experimentellem Design.

## 4.2 Strukturierte / Standardisierte Beobachtung

Beobachtung ist strukturiert / standardisiert, wenn

- Verhaltenseinheiten festgelegt
- Verhaltenseinheiten klassifiziert in Categoriesystem (erprobt, Anwender traininiert)

Bsp: Bales Interaction Process Analysis (IPA)

### 4.2.1 Beobachtungsfehler

#### 1. Fehler des Beobachters

- Wahrnehmungsfehler
  - Konsistenzeffekt (Konsistenz mit Gesamteindruck / vorherigem Eindruck. Bsp: Halo-Effekt) Abhilfe: Reihenfolge Permutieren
  - Einfluss vorangegangener Informationen (Konformitätseffekt, Erwartungseffekt, Konsistenzeffekt, Nachbarschaftseffekt) Abhilfe: Effekte erklären, Vermeidung (Keine Kommunikation zulassen)
  - Erwartungseffekt (Verhalten wird erwartungskonform interpretiert) Abhilfe: Effekte erklären, Beobachter kennen Hypothese nicht, Teilnehmende Beobachtung vermeiden
  - Observer drift (Systematische Veränderung durch Beobachtungsdauer)
    - \* Vergessen der Kriterien
    - \* Verschiebung der Gewichte
    - \* Zunehmende Ermüdung
    - \* Verbesserte Wahrnehmung
- Interpretationsfehler
  - Zentrale Tendenz (Vermeidung von Extremurteilen)
  - Persönliche Tendenz (Milde/Strenge, Ja/Nein, Kontrastbildung, Soziale Konformität) Abhilfe: Große Beobachterstichprobe, gute Schulung
- Erinnerungsfehler
  - Kapazitätsgrenzen (Kurzzeitgedächtnis auf 4-7 Einheiten je 15 Sekunden begrenzt) Abhilfe: Unmittelbare Kategorisierung / Protokollierung
  - Erinnerungsverzerrung / -selektion (Langzeitgedächtnis, rekonstruktiver Vorgang: Verzerrung der Information, Selektierung. Bsp: Etikettierungseffekt) Abhilfe: Unmittelbare Protokollierung

- Wiedergabefehler (Konformitätsdruck) Abhilfe: Wiedergabe unabhängig von anderen Beobachtern

## 2. Fehler der Beobachtung

- Reaktivität (Anderes Verhalten unter Beobachtung, Besonders stark bei Erwachsenen, aufdringlicher Beobachtung, bekannter Fragestellung) Abhilfe: Doppelblindversuch, Therapeutische Kontrollgruppe
- Beobachtungsbedingung (ist keine typische Situation für Verhalten)
- Probleme des Beobachtungssystems (zu viele / zu wenige Kategorien)

### 4.2.2 Kategoriesysteme

- Empirisch erprobt (Trennschärfe, Beobachtungsübereinstimmung)
- Strukturierte Kategorien (Hierarchisch strukturierte Kategorien)
- Registrierung von Beobachtungseinheiten (Phänomene genau beschrieben, voneinander abgegrenzt)
- Explizite Zuordnungsvorschriften (Definition der Kategorie, Beispiele, Regeln)

#### Segmentierung (Einheitenbildung)

- Formal
  - Zeitintervall (Zeittaktverfahren), Zeittakt i.d.R. 1 - 30 Sekunde
  - Person als Einheiten (Nur ein Ereignis pro Person registriert)
  - Syntaktisch definierte Einheiten (Wort, Satz)
- Semantisch / Psychologisch
  - Natürliche Einheiten nach inhaltlichen Gesichtspunkten
  - Bsp: IPA (Bedeutung der Handlung, meist zusätzlich Zeitpunkt) – Equilibriumstheorie (Gruppenintegration durch Aufgaben gestört) – Art der Aktion wird kategorisiert. Bedeutungswechsel bei Wechsel der Person, Äußerungen mit verschiedenen Bedeutungen. I.d.R. Jeder Satz ist eine Einheit

#### Auswertungsmöglichkeiten

- Interaktionsprofil (Häufigkeitsverteilung über Kategorien)
- Häufigkeit der Interaktionsrichtung (Wer-zu-wem-Matrix)
- Häufigkeit der Interaktionsrichtung pro Kategorie (Wer-zu-Wem-Matrix pro Kategorie)
- Häufigkeit der pro Kategorie und Zeiteinheit (Kategoriensequenzanalyse)

### 4.2.3 Reliabilität der Beobachtung: Beobachtungsübereinstimmung

Quantitativer Vergleich von Beobachtungsdaten mehrerer unabhängiger Beobachter (Kreuztabelle). Bei ungenügender Übereinstimmung: Weitere Trainings, da Reliabilität (und Validität) leidet.



## 5 Interview und Fragebogen

### 5.1 Beobachtung, Fragebogen oder Interview?

Beobachtung nicht möglich bei:

- Reaktivität
- Nicht spontan auftretendem Verhalten (Auslöser nicht gegeben)
- Verhalten, das sich der Beobachtung entzieht (Sexualverhalten, Drogen)

Verhalten muss erfragt werden. Befragter soll anonym bleiben  $\Rightarrow$  Interview / Fragebogen

Verhalten muss dafür bewusst sein. Erinnerungsfehler kaum vermeidbar.

Interview:

- Komplizierte Fragen (Nachfragen möglich)

Fragebogen:

- Noch nicht über Thema nachgedacht
- Antwortsituation peinlich
- Antworttendenz durch soziale Erwünschtheit

### 5.2 Arten und Ziele von Fragebogen

#### 5.2.1 Grad der Standardisierung

- Nichtstandardisiert (Fragen nicht ausformuliert, Fragereihenfolge und Antwortmöglichkeiten nicht festgelegt. Vgl. Interviewleitfaden) Ziel: Exploration eines Themas
- Teilstandardisiert (Fragen exakt ausformuliert, Antwortmöglichkeiten nicht festgelegt – Offenes Antwortformat)
- Vollstandardisiert (Alles ist genau festgelegt)

Ziel: Standardisierung zur Erhöhung von:

- Objektivität (Keine Interpretation der Antworten + Zuordnung zu Kategorie nötig, durch Vorgabe von Antwortmöglichkeiten)
- Reliabilität (Feste Antwortmöglichkeiten sorgen für ähnliche Ergebnisse, bei Wiederholung)
- Validität (Ausschluss von Missverständnissen der Fragen, durch vorgegebene Antworten)

### 5.2.2 Hauptanwendungsgebiete für Fragebogen

1. Diagnostischer Fragebogen (Erfassen von Persönlichkeitsmerkmalen / Persönlichkeitsdimensionen)
2. Demoskopischer Fragebogen (Faktenwissen, Meinungen, Einstellungen)

Bedingungen für Befragungsgegenstand:

- Bekannt
- Bewusst
- Sprachlich Beschreibbar

## 5.3 Schritte der Fragebogenkonstruktion

### 5.3.1 Inhaltliche Festlegung

Definition der zu erfassenden Verhaltensmerkmale durch

- Verhaltensbeobachtung
- Bekannte Symptome
- Aus Theorie abgeleitete Anhaltspunkte

Formulierung von Fragen  $\Rightarrow$  Operationalisierung

### 5.3.2 Festlegung des Fragentyps

Typ der Frage ist vom Inhalt abhängig. Richtiges Fragenformat abhängig von Funktion der Frage.

- Ziel der Frage: Funktionsfragen
  - Motivationsfragen (Eisbrecherfragen)
  - Kontrollfragen
    - \* Erhebungskontrollfragen (Zeit, Ort, Alter, Geschlecht, ...)
    - \* Auskunftskontrollfragen (Wiederholungsfragen zur Konsistenzermittlung)
  - Ablauforganisationsfragen
    - \* Filterfragen (Ausschluss von Personen)
    - \* Verzeigungsfragen (Unterschiedliche Frageabläufe)
  - Befragungstaktische Fragen
    - \* Ablenkungsfragen / Füllfragen (Verschleierung des eigentlichen Themas)
    - \* Pufferfragen (Ausschluss von Folgeeffekten, wie z.B. Halo-Effekt)
- Frageinhalte
  - Fakten
  - Wissen

- Bewertungen, Meinungen, Einstellungen
- Verhalten, Handlungen
- Motive
- Indirekte Fragen
  - Fragen durch die Hintertür / Fragen über Ersatzdimensionen (Nicht direkt fragen, sondern nach korrelierendem Merkmal. z.B. Jahrgang statt Alter)
  - Assoziationsfragen (Assoziierte Vorstellungen. Stellen sie sich vor ... Trifft das zu?)
  - Projektionsfragen (Hineinversetzen in andere. Was denkt... ?)
  - Dialogfragen (Meinung durch Schiedspruch. Wer hat recht?)
  - Hypothetische Situation (Wie würden Sie reagieren?)
  - Fragen mit Zitaten (Simmen Sie Zitat zu?)
- Form der Frage
  - Offene Frage (Frei Formulierte Antwort)
  - Geschlossene Frage (Antwortmöglichkeiten)

Fragetechniken bei unangenehmen Inhalten

- Möglichkeitsfragen (Könnte / Fänden)
- Jedermann-Ansatz (Alle... Und Sie?)
- Der Andere-Ansatz (Kennen Sie Leute? Was ist mit Ihnen?)
- Kinsey-Technik (Selbstverständlichung des peinlichen Inhalts. )

Standardisierung mit geschlossenen Fragen:

- Bekanntheit aller möglichen Antworten (Erschöpfende Antwortkategorien) nötig
- Optimierung der Rücklaufquote (Vermeidung von Offenen Fragen, da Beantwortung anspruchsvoll)
- Reproduktion (Offene Fragen besser. An was erinnern Sie sich?) / Wiedererkennen? (Geschlossene Frage. Einstufung auf Skala)

Arten geschlossener Fragen

- Alternativfragen (Ja / Nein)
- Selektivfragen (Multiple-Choice. Erhöhung der Komplexität des Fragebogens. Schlecht für Rücklaufquote)
- Ratingskalen (Merkmalsgröße / Bewertung. 5-7 Skalenpunkte üblich, 7-9 optimal. Rating durch verbale Benennung / Rating durch Ziffern)

### 5.3.3 Formulierung von Fragen und Antwortkategorien

- Eindeutige / Eindimensionale Fragen (Nur ein Inhaltsaspekt)
- Effekte extremer Formulierung: Neutralität (Kein beantwortungsvorteil einer Alternative). Aber: Extreme Antwortmöglichkeiten erhöhen Antwortwahrscheinlichkeit der anderen, nicht neutralen Antwortmöglichkeiten (Stimme uneingeschränkt zu / Lehne völlig ab)
- Reihenfolgeeffekte (Positionseffekt: Vorteil von Antwortalternativen an Anfang und Ende) Vermeidung durch permutierte Fragebögen (Split-ballot-Verfahren) = Randomisierung
- Rest- und Ausweichkategorie (Wenn zutreffende Antwort nicht unbedingt in Antwortmöglichkeiten enthalten). Nicht bei unangenehmen Themen verwenden, da verführerisch
- Vermeidung von Antworttendenzen (response set. Vermindern Validität)
  - Soziale Erwünschtheit (Vermeidung z.B. durch Jedermann-Ansatz)
  - Bejahungstendenz / Verneinungstendenz (Vermeidung durch Positive / Negative Fragen. Doppelte Verneinung fehleranfälliger)
  - Mittentendenz. Abhilfe: Keine extreme Formulierung, Verzicht auf Mitte, Erhöhung der Skalenpunkte, Mitte der Skala nicht auffällig
  - Tendenz zur vorschnellen Antwort (Antwort steht fest, bevor Frage fertig gelesen wurde) Abhilfe: Erst definieren, dann fragen.

### 5.3.4 Reihenfolge der Fragen

- Inhaltliche Zusammenhänge (Fragenblöcke mit logischem Aufbau. Diagnostisches Fragebögen: meist Zufallsreihenfolgen, Vergleich von Antworten erschweren, Inhalt Verschleiern)
- Minimierung von Reihenfolgeeffekten
  - Aktualisierungseffekt (Frage beeinflusst Beantwortung einer anderen, durch Bewusst machen von Sachverhalten) Abhilfe: Verwendung eines Trichters, Frageabfolge vom Allgemeinen zum Speziellen
  - Konsistenzeffekt (Verhalten muss mit Einstellung konsistent sein). Abhilfe: Erst nach Verhalten fragen, dann nach Einstellung
  - Redundanzeffekt (Befragter will sich nicht Wiederholen, wenn Antwort bereits bei anderer Frage gegeben)

### 5.3.5 Erprobung und Überarbeitung des Fragebogenentwurfs

- Prätest (Erste Erprobung eines Fragebogens, kleine Stichprobe, Interviewer anwesend). Zu klären:
  - Verständlichkeit (Nachfragen? Konsistenzanalyse: Antworten auf ähnliche Fragen sollen korrelieren)
  - Vollständigkeit der Antwortalternativen (Werden neue Alternativen benötigt?)

- Kontrolle von Reihenfolgeeffekten (Treten Reihenfolgeeffekte auf?)
- Bleiben Fragen unbeantwortet? (Warum?)
- Dauer der Durchführung (Zumutbar? Motivationsprobleme?)
- Itemselektion (Selektionsverfahren. Hauptsächlich: Trennschärfeanalyse)

Fragebogenauswertung und Trennschärfeanalyse

1. Wahl einer geeigneten Präteststichprobe (Hohe Varianz des Merkmals. Möglichst Normalverteilung)
2. Quantitative Auswertung der Antworten (Voraussetzung: Quantitative Dimension. Antworten bekommen Wert zugeordnet)
3. Ergebnismatrix
4. Trennschärfeanalyse:
  - Trennschärfe: Gute Diskriminierung aufgrund des Merkmals (Vergleich Frage, Gesamter Fragebogen)
  - Gute Trennschärfe: Hohe Korrelation mit Schätzung des Merkmals auf Basis des ganzen Fragebogens
  - Korrelation ist Trennschärfekoeffizient
  - Nur Fragen mit hohem Trennschärfekoeffizienten ( $>0,8$ ) im endgültigen Fragebogen berücksichtigen
5. Bestimmung der Reliabilität (Endgültiger Fragebogen, Stichprobe  $> 100$ ). Guter Wert  $> 0,9$ 
  - Split-Half-Reliabilität: Korrelation zwischen zwei Hälften des Fragebogens
  - Retest-Reliabilität: Zweimalige Anwendung an selber Stichprobe
  - Paralleltest-Reliabilität: Zwei Fragebogen zu selbem Merkmal (Paralleltest). Bearbeitet von selber Stichprobe
6. Validierung
  - Demoskopischer Fragebogen: i.d.R. nicht nötig
    - Vergleich mit anderen Verfahren zur Erfassung des selben Verhaltens (Kriteriumsvalidität)
    - Messung von Persönlichkeitsmerkmalen (Konstruktvalidität)
    - Korrelation mit Experimentaldaten (Konstruktvalidität)

## 6 Experimentelle Versuchsmethodik

### 6.1 Das Experiment

Beobachtung, Fragebogen, Interview erfassen nur ungerichtete Kovariation.

Hypothesentests für gerichtete Kovariation erfordern systematische Variation der UV  $\Rightarrow$  Experiment (Königsweg der Forschung).

Definitiorische Merkmale

- Willkürliche Herbeiführung des Phänomens
- Systematische Variation der UV
- Messung der AV
- Experimentelle Kontrolle (u.A. Konstante Bedingungen)

## 6.2 Variablentypen

Variablen sind

- Systematisch Veränderlich
- Veränderung erzeugt / registriert

Variablen Typen

- Unabhängige Variablen
- Abhängige Variablen (kausal abhängig von den UV)
  - Leistungsmaße (Operationalisiertes Leistungsverhalten: Reaktionszeit, Fehlermaße, Physiologische Parameter, wie Hautwiderstand, Pulsfrequenz, Elektroenzephalogramm – ereigniskorelierte Potentiale)
  - Subjektive Größen (Schätzurteile, Skalierung, Verbale Berichte – Introspektion wenig objektiv / reliabel)
- Störvariablen / Kontrollvariablen (Nicht Gegenstand der Hypothesenprüfung aber evtl. trotzdem Kausalfaktoren)

## 6.3 Prinzipien der Operationalisierung

### 6.3.1 Operationalisierung UVs

Überführung der theoretischen Begriffe in Beobachtungssprache

- Erzeugen / Nichterzeugen eines Effekts
- Experimentalgruppe / Kontrollgruppe
- Solomon-Four-Groups-Design

Abhängige / Unabhängige Stichproben

- Person ist eigene Kontrollgruppe (Abhängig)
- Kontrollgruppe unabhängig von Experimentalgruppe (Unabhängig)

Erzeugen einer Variation

- Instruktionsunterschiede (Unterschiedliche Angaben. Manipulation Check nötig – Ausreichend starke Unterschiede)
- Wechsel der Aufgabe (Schwierigkeitsgrad)

- Wechsel der Reizmuster (Wörter / Pseudowörter, Konsistente / Inkonsistente Angaben)
- Räumliche / Zeitliche Variation (Länge, Helligkeit, Kontrast, Zeitintervall ISI – Inter-Stimulus-Intervall, Räumliche Position)
- Verschiedene Zeitpunkte (Längsschnittstudie – Abhängige Stichproben)
- Vpn unterschiedlichen Alters (Querschnittstudie; Alter als UV. Nur Quasiexperiment, Korrelationsstudie)
- Vpn mit unterschiedlichen Merkmalen (Stichprobenselektion; Ähnlich Querschnittstudie)

### 6.3.2 Operationalisierung AVs

Kausal bedingt, beobachtbar / Messbar. Zeigen Kovariation mit UVs

- Objektive Indizes vor Subjektiven Größen (Auswählen besser als Bewertung)
- Konvergierende Operationen (Multiple Operationalisierung; Vermeidung von Reduzierung der Bedeutung ⇒ Mehrere AVs zur Überprüfung)

## 6.4 Probleme bei Messung von AVs

### 6.4.1 Neutrale Messung bei Selbst- und Fremdbeobachtung

Selbstbeobachtung ist nicht Wiederholbar. Keine neutrale Messung, da Messung Ergebnis beeinflusst. Doppelaufgabe: Vpn + Beobachter

### 6.4.2 Reliabilität von Messungen / Messgenauigkeit

Test-Retest / Paralleltest-Reliabilität

Probleme:

- Phänomen ist Problem
- Messinstrument ist Problem

Systematischer oder Zufallseinfluss möglich

Messfehler misst Variation der Messungen bei gleichen Merkmalen. Hoher Messfehler ⇒ Starker experimenteller Effekt wird unwahrscheinlich

## 6.5 Probleme der Validität eines Experiments

### 6.5.1 Experimentelle Kontrolle

Experimente zuverlässig und generalisierbar, wenn keine systematischen Effekte durch Störvariablen auftreten.

- Interne Validität (Störvariablen ausgeschlossen)
- Externe Validität (Generalisierbarkeit des Experiments – AVs, UVs, Stichprobe, Situation Repräsentativ?)

⇒ Experimentelle Kontrolle

Kontrolltechniken lassen Einfluss von Störvariablen auf AVs den Wert 0 annehmen (Unterschiede von Experimental- und Kontrollgruppe, jeweils Konstanter Einfluss der Störvariablen).

### 6.5.2 Sicherstellung interner Validität

Kofundierte Effekte (Störvariable + UV ergeben anderen Effekt, als UV)

- Bedingungskonsistenz (gleiche Rahmen- / Randbedingungen)
  - Situative Randbedingungen sind gleich (visuelle, auditive Reize)
  - Versuchsleitereffekte sind ausgeschlossen
  - Zeit ist systematisch kontrolliert oder gleich
- Ausbalancieren
  - Ausschluss von Reihenfolgeeffekten
  - Permutation der Versuchssequenzen; Alle möglichen Sequenzen gleich häufig
  - unvollständiges Ausbalancieren bei hoher Permutationsanzahl (Jede Bedingung zu jedem Messzeitpunkt gleich häufig)
- Randomisierung (Störvariable wirken nur zufällig und gleich verteilt; Nur bei großen Stichproben, Auch bei vielen Bedingungen möglich, Bei kleinen Stichproben → Ausbalancieren). Zufällige Zuordnung von:
  - Vpn
  - Reihenfolge
  - Messzeitpunkte bei Messwiederholung
- Parallelisierung und Messung kovariierender Variablen (Gleiche Versuchsgruppen hinsichtlich einer Variable)
  - Parallelisierung (n-Tupel mit nahezu gleicher Variable, zufällig auf n Gruppen aufgeteilt. Typische Variablen: Geschlecht, Alter, IQ)
  - Paarbildung (Paare mit nahezu gleichen Werten bei 2 Gruppen)

### 6.5.3 Sicherstellung externer Validität

Externe Validität: Bei gegeben Interner Validität → Generalisierung der Ergebnisse

- Repräsentativität der Operationalisierung (Generalisierung auf theorie Ebene möglich?)
- Repräsentativität der Stichprobe (Verallgemeinerung auf beliebige / bestimmte Population)
- Repräsentativität der experimentellen Situation (Ökologische Validität; Verallgemeinerung auf natürliche Umwelt?)



Sicherung externer Validität auf Seiten der

- UV (Operationalisierung deckt Theorie vollständig / wesentlichen Aspekte ab)
  - Werte der UV geeignet, um Effekte zu erzielen? (Manipulation Check)
  - Stufen der UV charakteristisch für natürliche Situation (Lohhausenstudie)
  - Andere Operationalisierung andere Wirkung?
- AV (Operationalisierung deckt Theorie vollständig / wesentliche Aspekte ab; Geschwindigkeits-Genauigkeitsausgleich)
- Stichprobenauswahl (Verallgemeinerung auf Population möglich?)
  - Zufallsstichproben + Randomisierung
  - Selektierte Stichproben (Wenn bestimmte Merkmale notwendig, Zufallsstichprobe aus dieser Gruppe)
  - Große Stichproben (Zufällige Abweichung kleiner)
  - Repräsentative Stichprobe (Demographische Abbildung der Bevölkerung)
- Situationsgestaltung (Ökonomische Validität)
  - Materiel & Aufgabe (Muss Realität entsprechen)
  - Räumliche Umwelt (Fotos vs. 3D Modelle)

#### 6.5.4 Labor- vs. Feldexperiment

Vorteile des Feldexperiments: Natürliche Umgebungsbedingungen, Vpn unwissend

- Ökologische Validität (Gegeben, Aber Situation ist auch nur Auswahl von möglichen Situationen)
- Reaktivität (Ausschluss durch unwissende Vpn)

Vorteile des Laborexperiments:

- Operationalisierbarkeit (Einweisung der Vpn, Weitere Messungen)
- Experimentelle Kontrolle (Umgebungsgeräusche kontrollieren)
- Herstellbarkeit der Phänomene (Optische Täuschungen)

Ökologische Validität  $\Rightarrow$  Feldexperiment Interne Validität  $\Rightarrow$  Laborexperiment

## 7 Versuchspläne

### 7.1 Experimentelle Versuchsplanung

#### 7.1.1 Ziele der Versuchsplanung

Versuchsplanung: Hypothesenprüfung bei maximaler interner und externer Validität

- Operationalisiere UV
- Operationalisiere AV
- $\Rightarrow$  Maximiere systematische Variabilität AV
- Kontrolliere Störvariablen
- Minimiere Messfehler

Experimentelles Design: Anordnung von Experimental- und Kontrollbedingungen Mindestens eine UV und AV

Quasiexperimentelles Design: prototypische Versuchsanordnung, Validität unbekannt. Häufig keine UV, Interne Validität nicht durch Kontrolltechniken sichergestellt

### 7.2 Quasiexperimentelle Versuchspläne und ihre Erweiterung zu experimentellen Versuchsplänen

Zu untersuchendes wird herbeigeführt

- Keine Systematische Variation einer UV
- Experimentelle Kontrolle nicht ausreichend

Befunde mehrdeutig  $\Rightarrow$  Mangelnde interne Validität.

#### 7.2.1 Einmalige Nacher-Messung an einer Stichprobe

Erklärung durch Hypothese nicht möglich, da kein Vergleich.

#### 7.2.2 Vorher-Nacher-Messung an einer Stichprobe

Keine Vergleichsgruppe vorhanden. Erklärung nicht eindeutig möglich.

Experiment: Kontrollgruppe nötig. Besser Solomon-Vier-Gruppen-Design.

#### 7.2.3 Statistischer Vergleich vorgegebener Gruppen

Natürlich existierende Gruppen. Vergleich nicht ohne weiteres möglich. Längsschnittstudie besser.

## 7.3 Experimentelle Designs

### 7.3.1 Randomisierte Versuchspläne

#### 1. Einfaktorielle Designs mit unabhängigen Stichproben

- Einfaktoriell (Effekt einer UV auf die AV)
- Faktor (Diskrete Stufen – Versuchsgruppen / Quantitative Abstufungen – Zeit)
- Unabhängig (Stufen sind verschiedene Versuchsgruppen; Verhalten wird nur unter einer Stufe beobachtet)
- vollständige Randomisierung (Ausschluss aller Störvariablen; Zufällige Zuteilung der Vpn auf die Stufen)

#### 2. Mehrfaktorielle Design mit unabhängigen Stichproben

- Mehrfaktoriell (Effekt mehrere UV auf die AV)
- Auswertung:
  - Haupteffekte der UV (Unterscheidungen der Mittelwerte)
  - Interaktion der UVs (Unterschiedliche Steigung der Auswertungsgeraden)



#### 3. Designs mit Messwiederholungsfaktoren (Unterschiede in Vpn ⇒ Selbe Vpn mehrmals messen; Ist eigene Kontrollgruppe)

- Ein- / Mehrfaktorielle Designs möglich