

Methodentraining Internetforschung  
HIIG, 31. Mai 2013 – Dr. Merja Mahrt

# **SURVEYS**

# Themen

- Grundlagen quantitativer Sozialforschung
  - Beispiel Befragung
  - Gütekriterien
  - Stichproben und Repräsentativität
- Befragungsverfahren
  - Online-Befragung
- Tipps zu Datenbereinigung, -management und Analysestrategien

# **GRUNDLAGEN QUANTITATIVER SOZIALFORSCHUNG**

# Grundideen

- Erkenntnisgewinn durch Empirie
- Quantifizierung sozialer Phänomene
  - Untersuchung großer Fallzahlen
  - Entwicklung standardisierter Instrumente (Skalen, Indizes)
  - Statistische Prüfung
- Befragung
  - zur Ermittlung von Fakten, Wissen, Meinungen, Einstellungen oder Bewertungen (Verhalten?)
  - Information wird im Verlauf einer sozialen Interaktion eingeholt
  - Quantitativ: Standardisierung von Frageformulierung, Antwortmöglichkeiten, Fragenreihenfolge, Verhalten des Interviewers

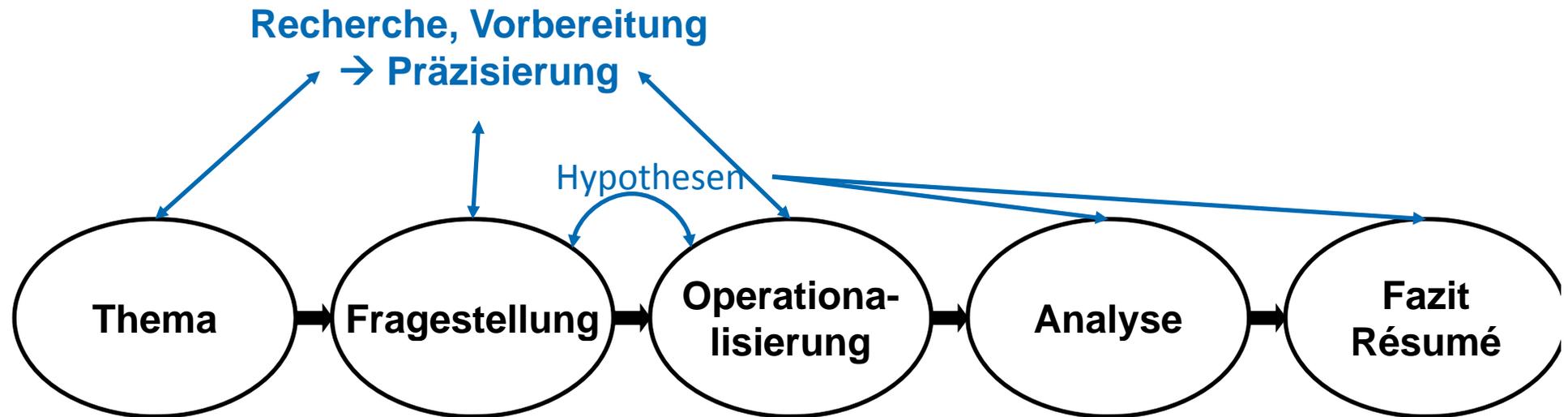
# Forschungsprozess



# Prinzipien

- Beantworten forschungsleitender Fragen  
oder
- Prüfen von Hypothesen
  - Ableitung aus existierender Literatur
  - Festlegung der zentralen Konstrukte
- FF1: Wie hängen Sozialkapital und Nutzung von sozialen Netzwerkseiten zusammen?
- H1: Nutzung sozialer Netzwerkseiten wirkt sich positiv auf Sozialkapital aus.

# Forschungsprozess



# Prinzipien

- Operationalisierung
  - Messbarmachung von theoretischen Begriffen, also die Überführung letzterer in beobachtbare Zustände in der Realität
  - Bei quantitativer Forschung i.d.R. Belegung von Zuständen in der Realität mit Zahlenwerten
- Messen
  - Zuordnung von Zahlen zu Objekten (in der Realität) nach bestimmten Regeln
  - Dabei sollen die einzelnen Messwerte zueinander in Beziehungen stehen, die den realen Beziehungen der Objekte entsprechen.

# Variable – Fall – Ausprägung

ZA3719\_1998 - SPSS Daten-Editor

Datei Bearbeiten Ansicht Daten Transformieren Analysieren Grafiken Extras Fenster Hilfe

1 : v1

	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10
1	3719	1	1	1	4	3	3	3	7	
2	3719	2	1	1	1	1	3	3	7	
3	3719	3	1	1	3	3	4	4	7	9
4	3719	4	1	1	2	2	3	3	1	
5	3719	5	1	1	3	3	3	4	5	
6	3719	6	1	1	2	2	3	3	7	
7	3719	7	1	1	2	2	3	4	7	
8	3719	9	1	1	5	3	3	3	7	
9	3719	10	1	1	2	2	3	3	6	
10	3719	11	1	1	5	5	5	5	7	
11	3719	12	1	1	5	4	5	4	7	
12	3719	13	1	1	5	5	5	5	7	
13	3719	14	1	1	3	3	2	3	2	
14	3719	15	1	1	4	4	3	3	6	
15	3719	16	1	1	2	2	3	1	1	
16	3719	17	1	1	5	4	3	3	7	
17	3719	18	1	1	1	2	3	3	7	
18	3719	19	1	1	4	3	2	2	7	
19	3719	20	1	1	9	9	9	9	7	
20	3719	21	1	1	5	3	5	3	7	
21	3719	22	1	1	4	3	4	3	6	
22	3719	23	1	1	5	5	5	5	6	
23	3719	24	1	1	4	3	4	3	7	
24	3719	27	1	1	4	2	8	3	7	

# Variable – Fall – Ausprägung

	Name	Typ	Spalte	Dezimal	Variablenlabel	Wertelabels	Fehlend	Sp	Ausricht	Meßniveau
1	v1	Nume	4	0	STUDIENNUMMER	{3719, ALLBUS COMPACT 1998}	Kein	8	Rechts	Ordinal
2	v2	Nume	4	0	IDENTIFIKATIONSNUMMER DES BEFRAGTEN	Kein	Kein	8	Rechts	Metrisch
3	v3	Nume	1	0	ERHEBUNGSGEBIET: WEST - OST	{1, ALTE BUNDESLAENDER}...	Kein	8	Rechts	Ordinal
4	v4	Nume	1	0	DEUTSCHE STAATSANGEHOERIGKEIT?	{1, JA}...	9	8	Rechts	Ordinal
5	v5	Nume	1	0	WIRTSCHAFTSLAGE IN DER BRD HEUTE	{1, SEHR GUT}...	7 - HI	8	Rechts	Ordinal
6	v6	Nume	1	0	WIRTSCHAFTSLAGE, BEFR. HEUTE	{1, SEHR GUT}...	7 - HI	8	Rechts	Ordinal
7	v7	Nume	1	0	WIRTSCHAFTSLAGE DER BRD IN 1 JAHR	{1, WESENTLICH BESSER}...	7 - HI	8	Rechts	Ordinal
8	v8	Nume	1	0	WIRTSCHAFTSLAGE, BEFR. IN 1 JAHR	{1, WESENTLICH BESSER}...	7 - HI	8	Rechts	Ordinal
9	v9	Nume	2	0	WICHTIGKEIT: EIGENE FAMILIE UND KINDER	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
10	v10	Nume	2	0	WICHTIGKEIT: BERUF UND ARBEIT	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
11	v11	Nume	2	0	WICHTIGKEIT: FREIZEIT UND ERHOLUNG	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
12	v12	Nume	2	0	WICHTIGKEIT: FREUNDE UND BEKANNTE	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
13	v13	Nume	2	0	WICHTIGKEIT: VERWANDTSCHAFT	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
14	v14	Nume	2	0	WICHTIGKEIT: RELIGION UND KIRCHE	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
15	v15	Nume	2	0	WICHTIGKEIT: POLITIK UND OEFFENTL. LEBEN	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
16	v16	Nume	2	0	WICHTIGKEIT VON NACHBARSCHAFT	{1, UNWICHTIG}...	97 - HI	8	Rechts	Ordinal
17	v17	Nume	1	0	FREIZEIT: BUECHER LESEN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
18	v18	Nume	1	0	FREIZEIT: ZEITSCHRIFTEN LESEN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
19	v19	Nume	1	0	FREIZEIT: LP, CD, CASSETTEN HOEREN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
20	v20	Nume	1	0	FREIZEIT: VIDEOKASSETTEN ANSCHAUEN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
21	v21	Nume	1	0	FREIZEIT: MIT DEM COMPUTER BESCHAEFTIGEN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
22	v22	Nume	1	0	FREIZEIT: INTERNET, ONLINE-DIENSTE NUTZEN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
23	v23	Nume	1	0	FREIZEIT: SICH PRIVAT WEITERBILDEN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
24	v24	Nume	1	0	FREIZEIT: EINFACH NICHTS TUN, FAULENZEN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal
25	v25	Nume	1	0	FREIZEIT: SPAZIERENGANGEN, WANDERN	{1, TAEGLICH}...	7 - HI	8	Rechts	Ordinal

# Skalenniveaus

Skalenart	Mögliche Aussagen	Beispiele
1. Nominalskala	Gleichheit Verschiedenheit	Farben
2. Ordinalskala	größer-kleiner Relationen	Bewertung von Kinofilmen
3. Intervallskala	Gleichheit von Differenzen	Temperatur in °C Jahreszahlen
4. Ratioskala	Gleichheit von Verhältnissen	tägliche Fernseh- nutzung in Minuten

} metrisch

# Skalenniveaus

- Je höher das Skalenniveau, um so mehr Auswertungsverfahren stehen dem Forscher zur Verfügung und um so höher ist auch der Informationsgehalt einer Variablen.

Bsp.:

„An wie vielen Tagen in der Woche nutzen Sie das Internet?“

0, 1, 2, 3, 4, 5, 6, 7

Ratioskala

0, 1-4, 5-7

Ordinalskala

nutzt das Internet ja / nein

Nominalskala

# Indizes

- Ein Index repräsentiert einen Merkmalsraum, in ihm sind mehrere Items rechnerisch zusammengefasst.

Bsp.: Persönlichkeitsstärke-Index

- „Gewöhnlich rechne ich bei dem, was ich mache, mit Erfolg.“
- „Ich bin selten unsicher, wie ich mich verhalten soll.“
- „Ich übernehme gern Verantwortung.“
- „Ich übernehme bei gemeinsamen Unternehmungen gern die Führung.“
- „Es macht mir Spaß, andere Menschen von meiner Meinung zu überzeugen.“
- „Ich merke öfter, dass sich andere nach mir richten.“ ...

# Skalierungsverfahren

## Likert-Skala

- „Auf einer Skala von 1 bis 5, wobei 1 bedeutet „Stimme voll und ganz zu“ und 5 „Stimme überhaupt nicht zu“ wie stehen Sie zu einer Altersbeschränkung für Film X?“

## Semantisches Differential

Film X ist...

spannend	1	2	3	4	5	langweilig
lustig	1	2	3	4	5	traurig
gewalthaltig	1	2	3	4	5	friedlich

# Formen der Befragung

Standardisierung	Kommunikationssituation			schriftliche Befragung (auch online)
	persönliches face-to-face Interview		Telefonbefragung	
	Einzelperson	Gruppe		
nicht	narratives Interview	freie Gruppendiskussion	(Recherchegespräch)	Aufsatz
teil	Leitfadeninterview	Metaplandiskussion		Tagebuchbefragung
voll	standardisierte Umfrage	(Gruppeninterview)	Telefonumfrage (CATI)	<b>schriftliche Befragung zum Selbstausfüllen („Survey“)</b>

# Alternativen zur Befragung

- **Inhaltsanalyse:** Untersuchung von Dokumenten
- **Beobachtung:** Untersuchung von Verhalten oder dessen Spuren

## **Mögliche Vorteile gegenüber Befragung**

- „Natürlichkeit“
- Nicht-Reaktivität: Datenerhebung beeinflusst den Untersuchungsgegenstand (meist) nicht
- u.U. sehr große Fallzahlen möglich

# Online-Alternativen zur Befragung

## **Inhaltsanalyse**

- Analyse von Blogposts, Forenbeiträgen, Webseiteninhalten...
- ... die nicht speziell angefertigt wurden

## **Beobachtung**

- Internetautnutzungsverhalten hinterlässt Spuren (Logfiles, „Likes“, „Retweets“, ...)

# **GÜTEKRITERIEN (UND WIE MAN SIE EINHÄLT)**

# Gütekriterien

## **Objektivität**

- Ein (quantitativer) Messvorgang sollte *objektiv*, also vom Versuchsleiter unabhängig sein.

## **Reliabilität**

- Ein Messinstrument soll Werte präzise und stabil, d. h. bei wiederholter Messung reproduzierbar, erfassen.
- Objektivität und Reliabilität sind notwendige Minimalanforderungen an ein Messinstrument.

# Gütekriterien

## Validität

- Ein Messinstrument ist *valide*, wenn es tatsächlich das misst, was es messen soll.
- Ein möglichst valides Messinstrument ist damit das Hauptziel jeder Erhebung.
- Allerdings lässt sich Validität nicht so einfach messen wie Reliabilität. Es werden immer Vergleichsmessungen oder -werte benötigt, deren Auswahl fraglich ist oder die, wenn sie als Maßstab gelten, den Sinn einer erneuten Messung in Frage stellen können.

# Antwortverzerrungen



# „Gebote“ zur Frageformulierung

- Einfache und unzweideutige Begriffe verwenden
- Lange und komplexe Fragen vermeiden
- Hypothetische Fragen vermeiden
- Doppelte Stimuli und doppelte Verneinungen vermeiden
- Unterstellungen und suggestive Fragen vermeiden
- Nur nach Informationen fragen, über die die Befragten auch tatsächlich verfügen können
- Eindeutige zeitliche Bezüge verwenden
- Erschöpfende und disjunkte Antwortkategorien verwenden (evtl. „weiß nicht“ vorsehen)
- Möglichst vermeiden, dass der Kontext einer Frage Auswirkungen auf die Beantwortung der Frage hat

# Offene vs. geschlossene Fragen

- Geschlossene Fragen geben alle Antwortoptionen vor
  - Vorteil: Standardisierung ermöglicht schnelle Auswertung
  - Nachteil: Antwortkategorien müssen zu den Vorstellungen der Befragten passen
- Offene Fragen lassen alle Arten von Antworten zu
  - Vorteil: Angaben werden nur vom Befragten gesteuert
  - Nachteil: für statistische Auswertung u.U. sehr hoher Codierungsaufwand
- Teilstandardisierung ist möglich
  - nach einer Reihe von vorgegebenen Kategorien folgt „sonstiges“, „außerdem finde ich noch wichtig“ ... sowie ein Freifeld

# Reihenfolge-/Kontexteffekt

- Drei Gruppen werden nach ihrer Einstellung zur CDU befragt (Skala von 0 bis 10)

Gruppe	A	B	C
Bewertung der CDU	3,4	5,2	6,5
Vorfrage	Wissen Sie zufällig, welches Amt R. v. W. ausübt, das ihn außerhalb des Parteigeschehens stellt?	(ohne)	Wissen Sie zufällig, welcher Partei R. v. W. seit mehr als 20 Jahren angehört?

# Skaleneffekt

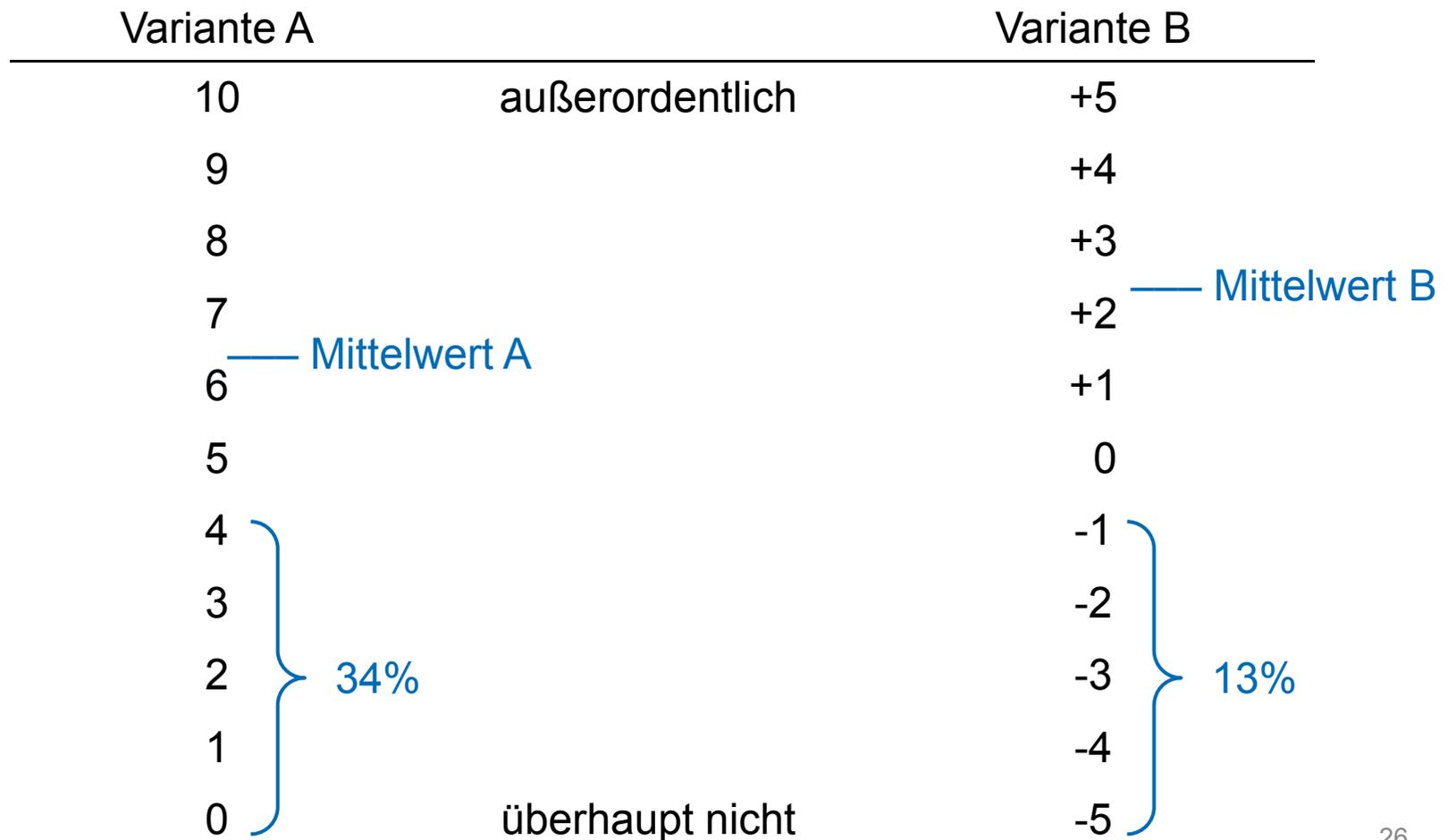
- Fernsehkonsum wird mittels unterschiedlicher Antwortvorgaben erhoben

Variante A	%	Variante B	%
bis 30 min	7		
bis 60 min	18		
bis 90 min	27		
bis 120 min	15		
bis 150 min	18	bis 150 min	63
mehr als 150 min	16	bis 180 min	23
		bis 210 min	8
		bis 240 min	5
		bis 270 min	2
		mehr als 270 min	0

Häder (2006), S. 15

# Noch ein Skaleneffekt

- Frage nach dem eigenen Erfolg im Leben



# Formulierungseffekt

Glauben Sie, dass die USA öffentliche Angriffe auf die Demokratie verbieten sollte?

Glauben Sie, dass die USA öffentliche Angriffe auf die Demokratie nicht erlauben sollte?

---

Ja

54%

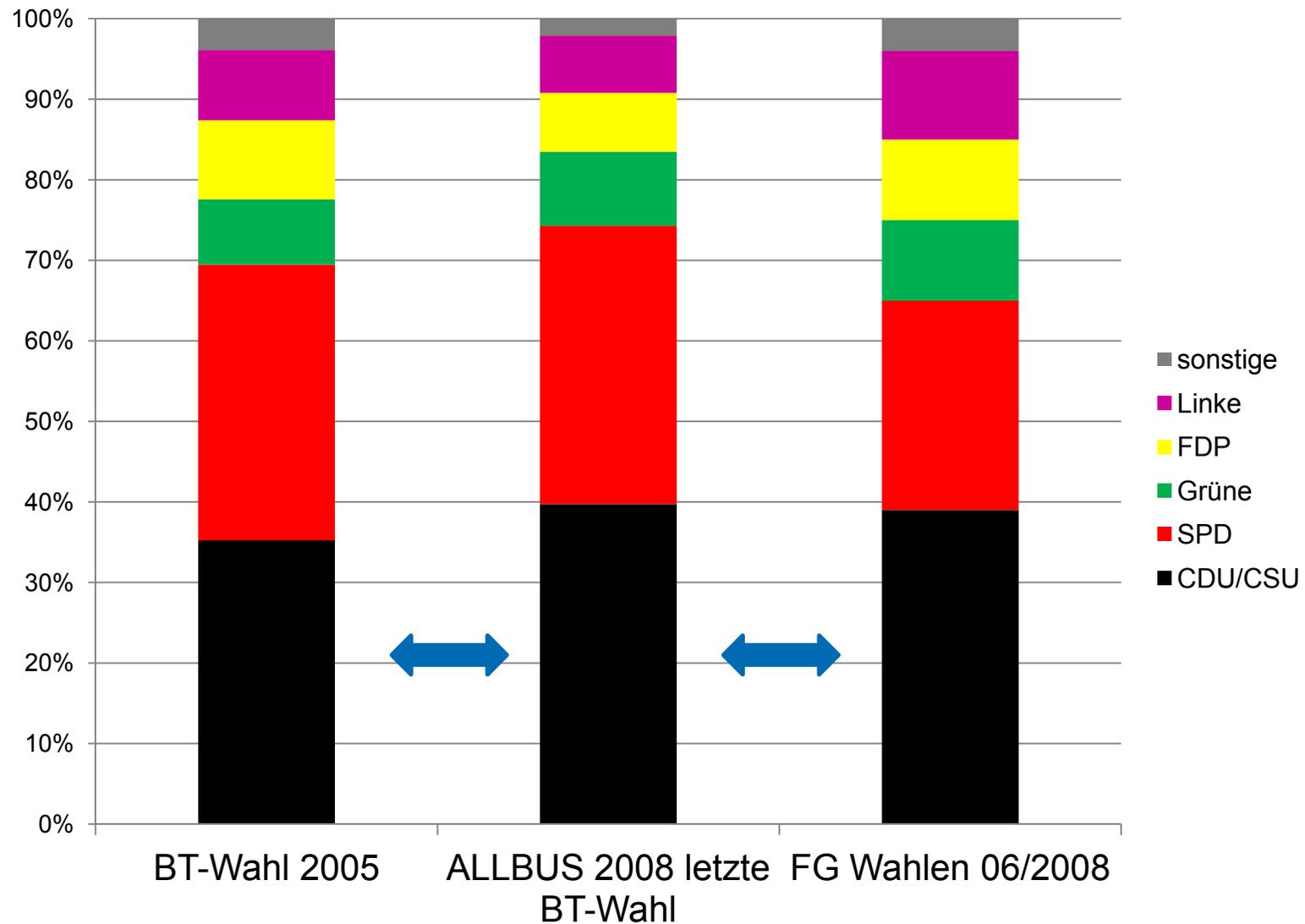
75%

Nein

45%

25%

# Beispiel – Wahl vs. Sonntagsfrage



# „Gebote“ zur Fragebogendramaturgie

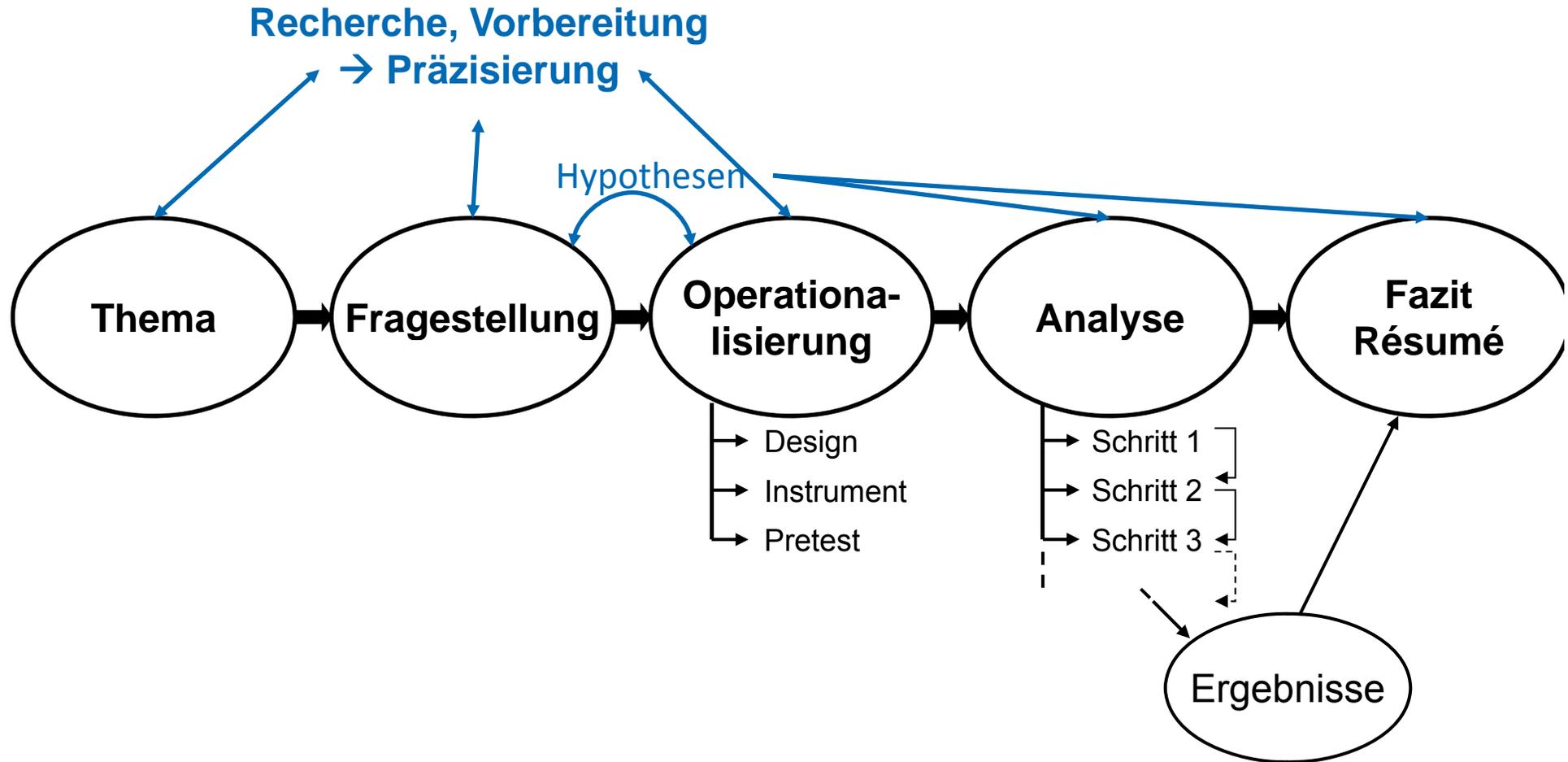
- Einführende Worte, u.U. Hinweis zu Anonymität, (Nicht-) Weitergabe der Daten, Auftraggeber, Ziele, Kontaktmöglichkeit...
- Einfach zu beantwortende, unverfängliche Fragen zu Beginn, idealerweise „positive“ Fragen
- Überleitungen zwischen Themenblöcken
  - evtl. Trichterform: vom Allgemeinen zum Speziellen
- Persönliche Daten eher weiter hinten (Alter, Einkommen...)
- Danksagung am Ende

# Pretest

- Erprobung des Instruments vor der Datenerhebung
- Umfang möglichst entsprechend 10% der geplanten Fallzahl, ähnlicher Personenkreis
- Klärung von Missverständnissen
- Überprüfung von Filtern
- Überprüfung von Skalen
- ggf. Sammeln von offenen Antworten zur Bildung von Item-Batterien in der Hauptuntersuchung
- ggf. Erstellen von Probeauswertungen zur Überprüfung der Vollständigkeit

# **STICHPROBEN UND REPRÄSENTATIVITÄT**

# Forschungsprozess



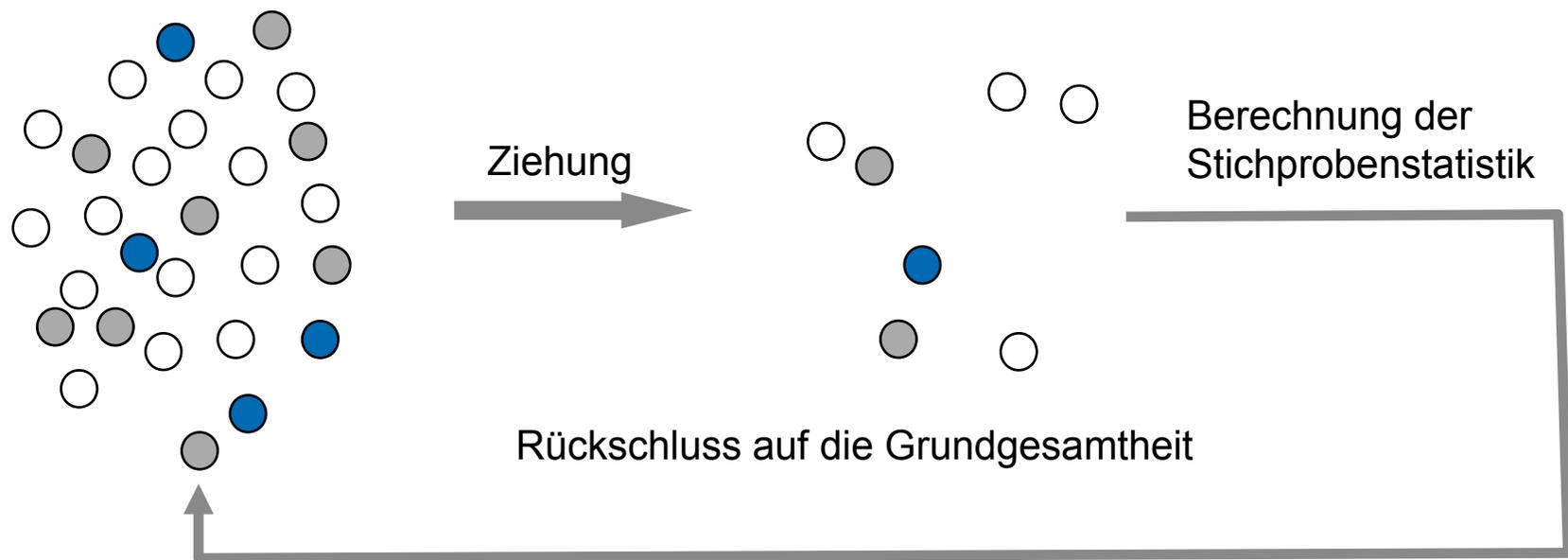
# Stichproben und Repräsentativität

Folgende Fragen stellen sich den Forschenden:

- Ist die Grundgesamtheit bekannt?
- Wie groß ist sie?
- Welches Vorwissen gibt es über die Grundgesamtheit?
- Auf welche Weise kann eine Auswahl aus dieser Grundgesamtheit vorgenommen werden?

# Stichproben und Repräsentativität

- Ziel einer Stichprobe ist die möglichst strukturgleiche Abbildung der Grundgesamtheit. Diese ist Grundlage für den späteren Repräsentativitätsschluss von der Stichprobe auf die Grundgesamtheit.



# Stichprobenarten und Repräsentativität

Wenn Repräsentativität angestrebt wird

- echte Zufallsstichprobe
- Quotenstichprobe

Keine repräsentativen Aussagen erlauben

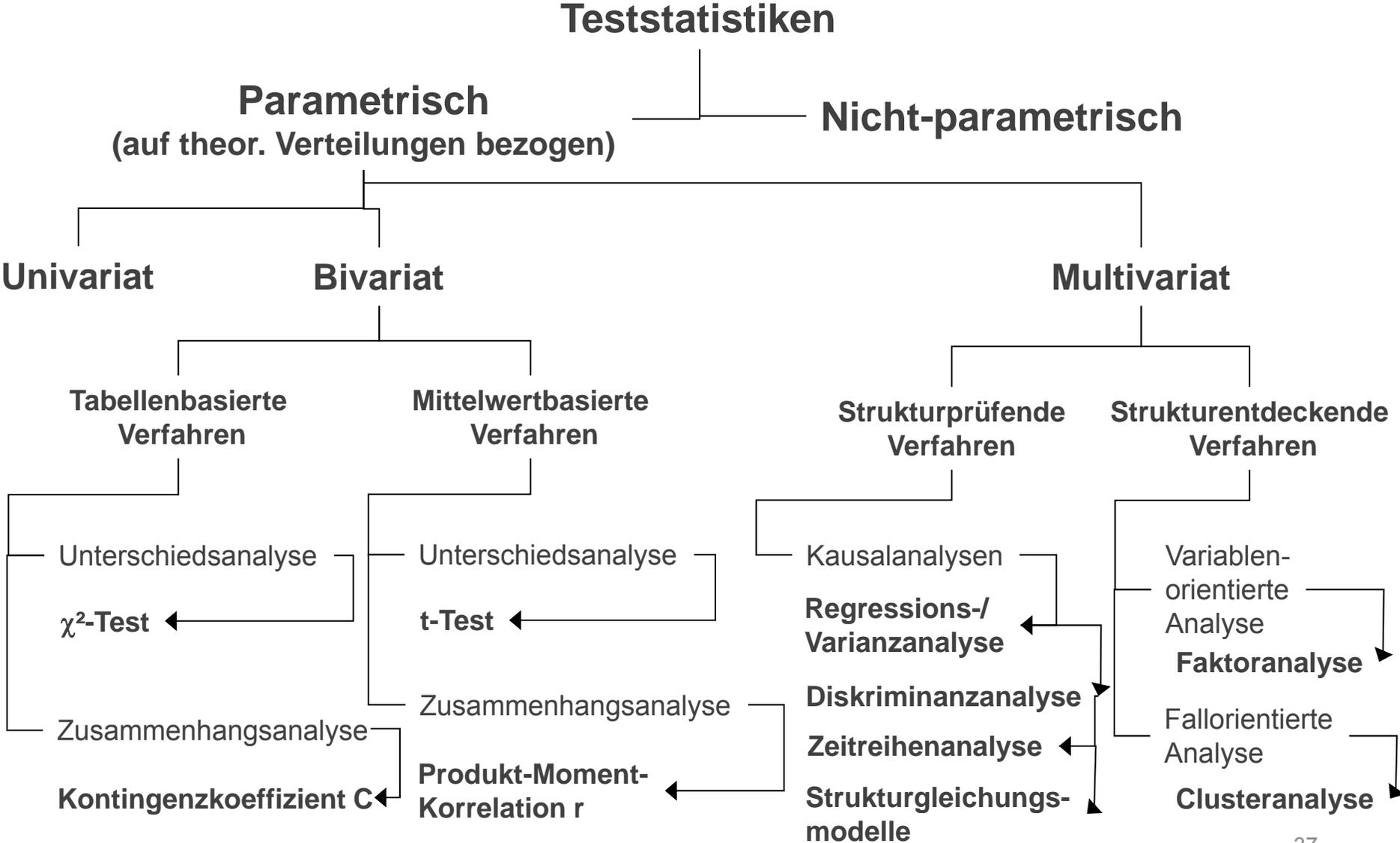
- Convenience samples

# Statistischer Umgang mit Stichproben

- Schätzt man einen unbekanntem Parameter der Grundgesamtheit (z.B. Mittelwert  $\mu$ ) mit Hilfe einer Stichprobenfunktion (z.B. Stichprobenmittelwert  $\bar{x}$ ), so *trifft* man den Parameter *niemals genau* (Problem der Punktschätzung).
- Mit Hilfe statistischer Verteilungen lässt sich jedoch eine Aussage darüber machen, in welchem Bereich sich der unbekanntem Parameter *mit einer vorgegebenen Wahrscheinlichkeit* (Sicherheitsniveau) befindet (**Intervallschätzung**).

$$\bar{x} \pm 1,96 \cdot \frac{s}{\sqrt{n}} \qquad p \pm 1,96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$$

# Einige statistische Möglichkeiten...



# **BESONDERHEITEN VON ONLINE-BEFRAGUNGEN**

# Datenerhebung online: Vorteile

- schnell
- kostengünstig
- automatisierte / schnelle / einfache Auswertung
- zeitlich / örtlich unabhängig
- Multimedialität ist möglich
- komplexe Filterführungen sind möglich

s. Welker & Matzat in Jakob et al. (2009)

# Datenerhebung online: Nachteile

- geringe Teilnahmebereitschaft
- nur internetaffine Menschen werden erreicht
- Repräsentativität fraglich
- Identität der Befragten nicht überprüfbar / Angaben können nicht extern kontrolliert werden
- geringe Aufmerksamkeit
- Abbruch der Befragung ist einfach möglich
- niedrige Rücklaufquote

s. Welker & Matzat in Jakob et al. (2009)

# Fragebogengestaltung online

- Nutzerfreundlichkeit
  - Texte eher kurz
  - Fragen klar hervorheben (Kontrast, größere Schrift...)
  - Fortschrittsanzeige
  - bei langen Fragebögen: Wiedereintrittsmöglichkeit bieten
- Zugänglichkeit für verschiedene Browser
- Layout
  - an schriftliche Fragebögen angelehnt
  - Farben nur sehr begrenzt einsetzen
  - möglichst kein Scrollen, ganze Seite sichtbar
  - möglichst keine Dropdown-Menüs

s. Maurer & Jandura in Jakob et al. (2009)

# Stichprobenziehung online: Typen

- offener Zugang: Link auf Webseite
- Freiwilligen-Panels (eher Marktforschung)
- Intercept-Befragungen: z.B. Zufallsstichprobe aus Besuchern einer Webseite
- Stichproben aus Listen: z.B. Mitglieder einer Organisation mit Verzeichnis
- Mixed-Mode-Befragung: Stichprobe „klassisch“ gezogen, Befragte können Version des Fragebogens wählen
- Access-Panels: oft repräsentativ für Internetnutzer eines Landes
- offline rekrutierte Panels der Gesamtbevölkerung

s. Baur & Florian in Jakob et al. (2009)

# Besonderheiten von Internetnutzern

- deutlich jünger als Bevölkerungsdurchschnitt
- Bessergebildete nutzen häufiger; andere Interessen
- Geschlechtsunterschiede gleichen sich an, Männer nutzen aber nach wie vor viele Dienste häufiger
- Berufstätige nutzen häufiger
- Besserverdienende nutzen häufiger
- in HH mit Kindern häufiger Internetnutzung

s. Baur & Florian in Jakob et al. (2009)

# Erhöhung der Rücklaufquote

- Befragte vorab kontaktieren
- Befragte persönlich ansprechen
- zunächst inaktive Zielpersonen mehrfach ansprechen
- Fragebogen übersichtlich, abwechslungsreich, nicht zu lang gestalten
- Incentives anbieten
  - (materielle) Incentives steigern die Teilnahmebereitschaft
  - Incentives sind individuell unterschiedlich attraktiv
  - Incentive darf nicht alleiniger Grund für die Teilnahme sein
  - ggf. über Cookies oder IDs mehrfache Teilnahme ausschließen

s. Maurer & Jandura in Jakob et al. (2009); Welker et al. (2005)

# Nach der Erhebung

- Auch online erhobene Daten müssen bereinigt und auf ihre Qualität sowie Plausibilität überprüft werden.
  - Vor allem: Sind Skalenwerte in den Daten richtig übernommen?
  - Haben Befragte (unerwünscht) mehrfach teilgenommen?
  - Gibt es auffällige Ausreißer? Wenn ja: Sind deren Antworten insgesamt plausibel?
- falls keine Korrektur möglich: problematische Fälle oder Variablen löschen

**NACH DER ERHEBUNG:  
DATENBEREINIGUNG & ANALYSE**

# Unbedingt...

1. Eine Version der Rohdaten aufbewahren.
2. Daten neu abspeichern, bevor sie verändert werden (= bevor Einträge in der Datenmatrix so verändert werden, dass es nicht rückgängig gemacht werden kann).
3. Dokumentation der Datenaufbereitung und Analyse lieber ausführlicher als kurz und kryptisch halten.
  1. Ein Log führen, z.B. durch Kommentare in Syntaxdateien:  
\*Modell mit Wahlbeteiligung als AV, ohne Mediennutzung.
  2. Für Dateien, Variablen, Labels... möglichst inhaltliche, deutliche Bezeichnungen wählen (Was war noch mal F3? In welcher Datei ist eine Analyse: Output1.spv, Output2.spv...?)  
Besser z.B.: Korrelationen\_Wahlbeteiligung\_2011-05-12.spv

# Unbedingt...

4. Umcodierungen, Neuberechnungen etc. immer prüfen.
5. Daten gründlich bereinigen.
6. Bei Sekundäranalysen: Große Datensätze zuschneiden.
  1. Fälle auswählen (nur Einwohner eines Landes, nur eine Panelwelle...)  
select if (welle=1).  
execute.  
Achtung: Damit werden die anderen Fälle gelöscht!
  2. Variablen auswählen  
save outfile='Dateipfad + Name' /keep v1 v2 v3.  
Eine Datei überschreiben oder eine neue speichern.
7. Eine ID-Variable „aufheben“, um ggf. später andere Variablen eindeutig zuführen zu können.

## Auch noch hilfreich...

1. Syntax benutzen.
2. Eine Grundauszählung machen, sich mit den Daten und Verteilungen vertraut machen.
3. Prüfen, ob man seine Daten gewichten sollte.
4. Zwischenspeichern.
5. Evtl. zusätzliches Speicherformat wählen, das man auch ohne Statistiksoftware öffnen kann.
6. Alle Variablen so drehen, dass ein hoher Wert einer (inhaltlich) starken Ausprägung entspricht.

## Auch noch hilfreich...

7. Zwischendurch Ergebnisse und Interpretationen festhalten, sich nicht dem Analyse-Flow hingeben.
8. Kein „Model Finding Business“, also so lange am Modell rumschrauben, bis  $R^2$  0,1 Punkte höher ist o.ä. Theorie und Forschungsinteresse leiten die Analyse!

# Analysestrategien

- Planvolles Vorgehen
  1. Daten bereinigen
  2. Zentrale Variablen deskriptiv auszählen, evtl. Korrelationen anschauen
  3. Evtl. Verdichtung von Variablen (z.B. Indizes)
  4. Evtl. „einfache“ Hypothesentests
  5. Multivariate Verfahren
  6. Modellgüte prüfen, evtl. Alternativmodelle vergleichen
  
- VORHER entscheiden: Was brauche ich?

# Das richtige Verfahren finden

- Wonach fragen Hypothesen?
  - Unterschiede, Zusammenhänge, Kausalbeziehungen?
  - Uni-, bi- oder multivariat?
  - Längsschnitt oder Querschnitt?
- Welche Skalenniveaus liegen vor?
  - Metrisch oder nicht metrisch?
- Was kenne und kann ich?
  - Bestehende Kenntnisse
  - Mögliche Betreuung
  - Vorhandene Software

# Literatur – Datenerhebung

- Diekmann, A. (2010). *Empirische Sozialforschung* (12. Aufl.). Reinbek: Rowohlt.
- Häder, M. (2006). *Empirische Sozialforschung. Eine Einführung*. Wiesbaden: VS Verlag.
- Jackob, N., Schoen, H., & Zerback, T. (Hrsg.). (2009). *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung*. Wiesbaden: VS Verlag.
- Schnell, R., Hill, P. B., & Esser, E. (2008). *Methoden der empirischen Sozialforschung* (8. Aufl.). München: Oldenbourg.
- Scholl, A. (2009). *Die Befragung. Sozialwissenschaftliche Methode und kommunikationswissenschaftliche Anwendung* (2. Aufl.). Konstanz: UVK.
- Vogt, W. P., Gardner, D. C., & Haeffele, L. M. (2012). *When to use what research design*. New York, NY: Guilford.
- Welker, M., Werner, A., & Scholz, J. (2005). *Online-Research. Markt- und Sozialforschung mit dem Internet*. Heidelberg: dpunkt.

# Literatur – Statistik

- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2008). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (12. Aufl.). Berlin: Springer.
- Bortz, J. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Berlin: Springer.
- Brosius, F. (2008). *SPSS 16*. Bonn: mitp.
- Miles, J., & Shevlin, M. (2001). *Applying regression and correlation. A guide for students and researchers*. London: Sage.
- Müller-Benedict, V. (2007). *Grundkurs Statistik in den Sozialwissenschaften* (4. Aufl.). Wiesbaden: VS Verlag.
- Wonnacott, T. H., & Wonnacott, R. J. (1990). *Introductory statistics for business and economics* (4. Aufl.). New York: Wiley.
- Zöfel, P. (2001). *Statistik verstehen. Ein Begleitbuch zur computergestützten Anwendung*. München: Addison-Wesley.

Vielen Dank

Dr. Merja Mahrt

Heinrich-Heine-Universität Düsseldorf  
mahrt@hhu.de