

Prof. Dr. Hendrik Godbersen

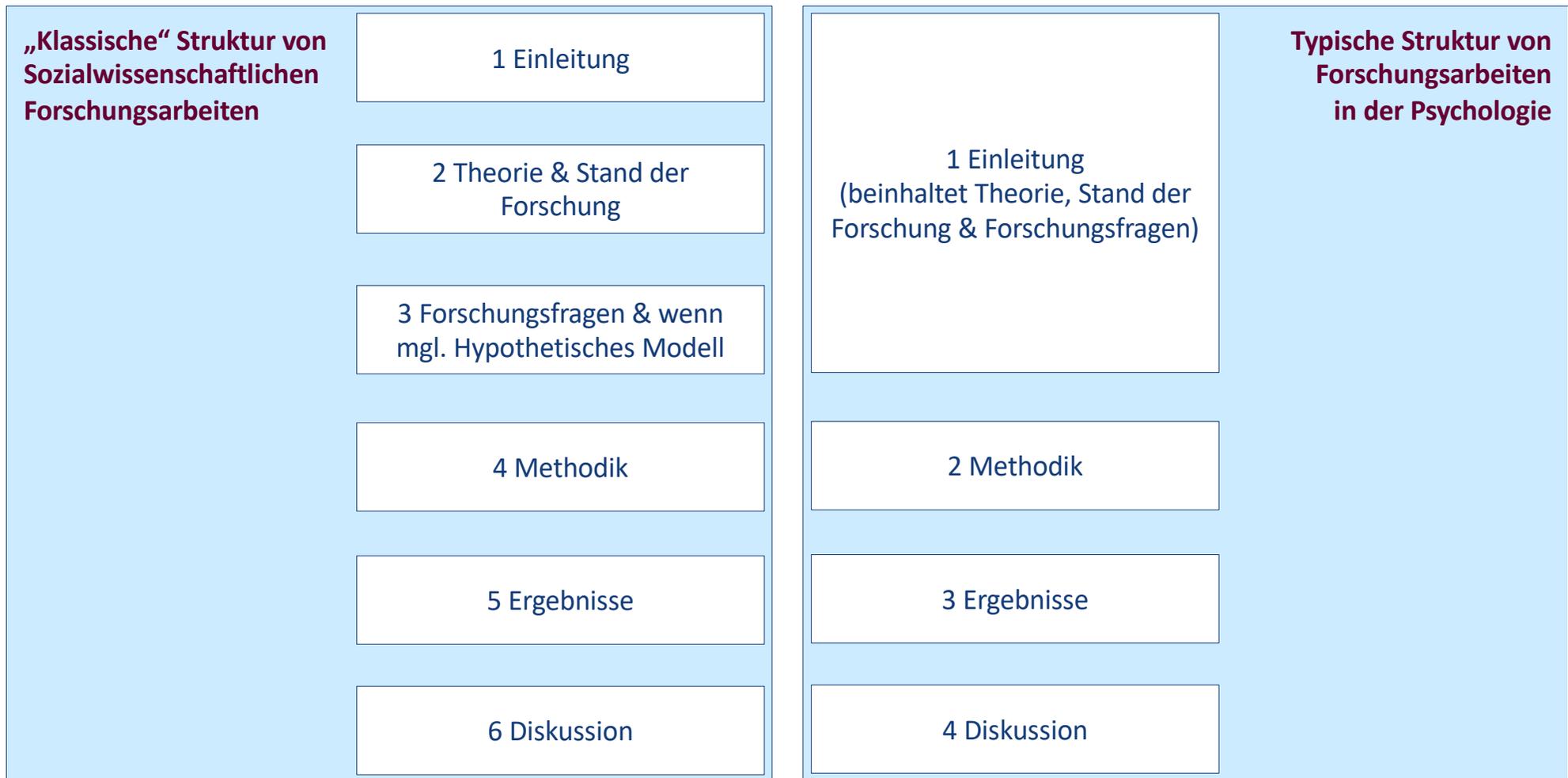
Crash-Kurs

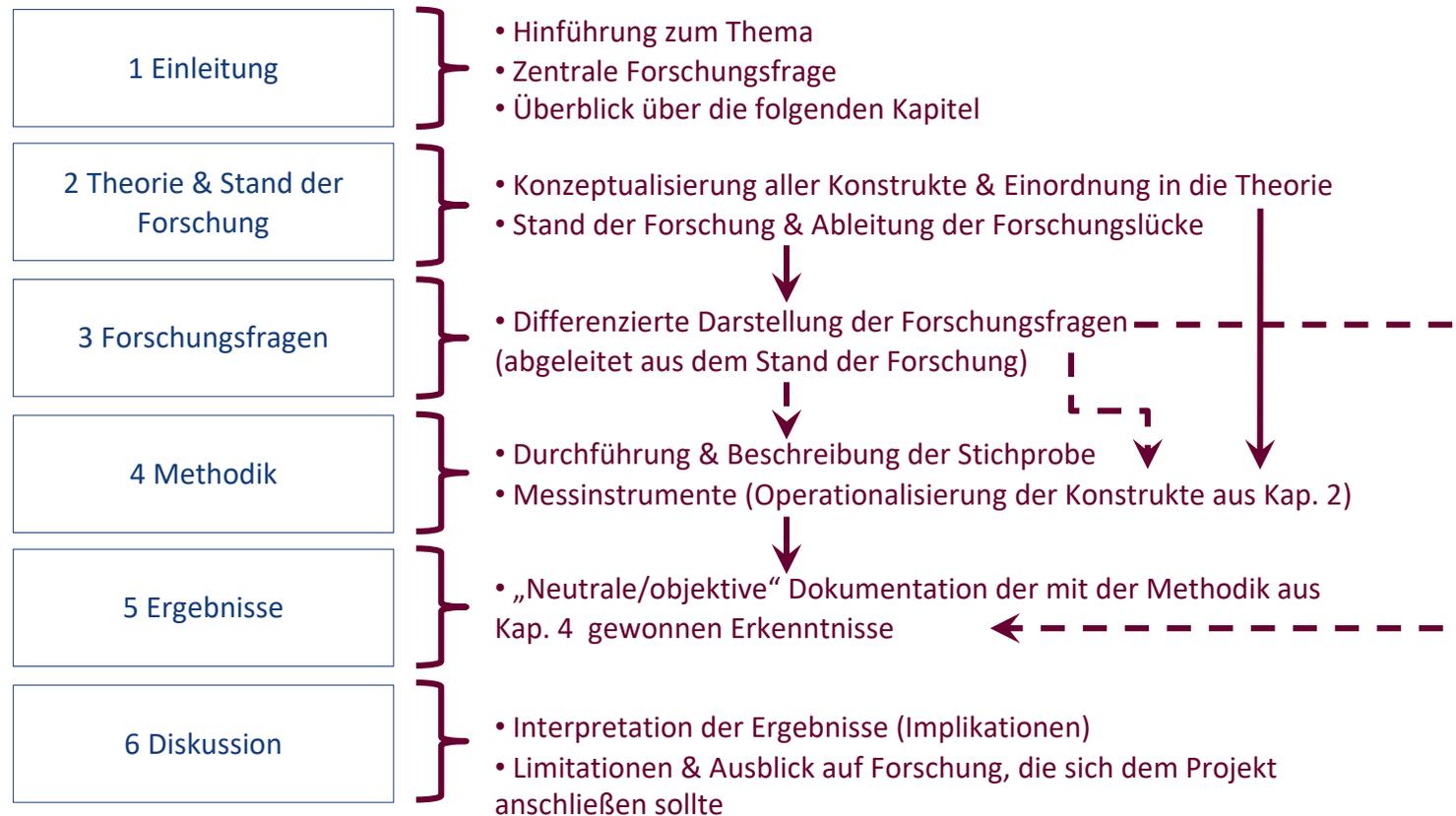
Quantitative Forschungsarbeiten

- (1) Es müssen **neue Erkenntnisse** gewonnen werden.
- (2) Es müssen **generalisierbare Ergebnisse** gewonnen werden.
- (3) Die neuen Erkenntnisse müssen mit einer **nachvollziehbaren und überprüfbaren Methodik** gewonnen werden.
- (4) Die Ableitungen müssen **zwingend logisch** sein.
- (5) Es muss eine **erschöpfende Behandlung** des Themas erfolgen.



„Klassische“ Struktur von sozialwissenschaftlichen Arbeiten





Determinanten von Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy

1 Einleitung

2 Theoretischer Hintergrund

2.1 Kundenzufriedenheit und Kundenbindung

2.2 Sharing Economy und Online-Plattformen

2.3 Marktplatz-Akzeptanzmodell nach Kollmann

2.4 Weitere Determinanten von Kundenzufriedenheit und Kundenbindung

3 Hypothetisches Modell und Forschungsfragen

4 Methode

4.1 Untersuchungsdesign

4.2 Messinstrumente

5 Ergebnisse

6 Diskussion

Quelle: Godbersen, H., Brühl, R. (2019): Determinanten von Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy, in: transfer – Zeitschrift für Kommunikation und Markenmanagement, Jg. 65, Nr. 4, S. 6-17

Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

- Zweck der Einleitung
 - Der Leser soll einen Überblick darüber erhalten, um was es in der Forschungsarbeit geht und was ihn in den folgenden Abschnitten erwartet.
- Inhalt
 - Kurze Hinführung zur vorliegenden Forschung (~„Aufhänger der Arbeit“)
 - Zielstellung der Arbeit / zentrale Forschungsfrage(n)
 - Überblick über die folgenden Abschnitte
- Empfehlung:
 - I.d.R. sollte die Einleitung am Ende des Forschungsprojektes geschrieben werden.

Quelle: Godbersen, H., Brühl, R. (2019): Determinanten von Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy, in: transfer – Zeitschrift für Kommunikation und Markenmanagement, Jg. 65, Nr. 4, S. 6-17

1 Einleitung

Die Sharing Economy ist ein Wirtschaftsbe- reich, der in den letzten Jahren stark gewach- sen ist. Online-Plattformen wie Ebay, Airbnb oder Uber sind wesentliche Akteure der Wirtschaft geworden. Vor diesem Hin- tergrund wird im vorliegenden Beitrag untersucht, welche Determinanten einen Einfluss auf die Kundenzufriedenheit mit und die Kundenbindung zu derartigen Online-Plattformen haben. Zu diesem Zweck werden im nächsten Abschnitt die theoretischen Grundlagen behandelt, um auf dieser Basis ein hypothetisches Modell der

Wirkzusammenhänge und die Forschungs- fragen zu entwickeln. Daran anschließend wird die Methodik der empirischen Unter- suchung erläutert sowie die Ergebnisse dar- gestellt und diskutiert.

2 Theoretischer Hintergrund

Im folgenden Abschnitt werden die Kunden- zufriedenheit und Kundenbindung sowie die Sharing Economy und Online-Plattformen eingeführt. Auf dieser Grundlage werden ausgehend vom Marktplatz-Akzeptanzmo- dell nach Kollmann (2001) die Determinan-

ten von Kundenzufriedenheit und Kunden- bindung konzeptualisiert.

2.1 Kundenzufriedenheit und Kundenbindung

Die Kundenzufriedenheit und die Kunden- bindung sind zentrale Begriffe im Kunden- beziehungsmanagement. So wird die Kun- denzufriedenheit als notwendige jedoch nicht hinreichende Bedingung für den öko- nomischen Erfolg von Unternehmen gese- hen (u.a. Bruhn, 2011). In der Literatur hat sich als Erklärungsmodell weitestgehend das

Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

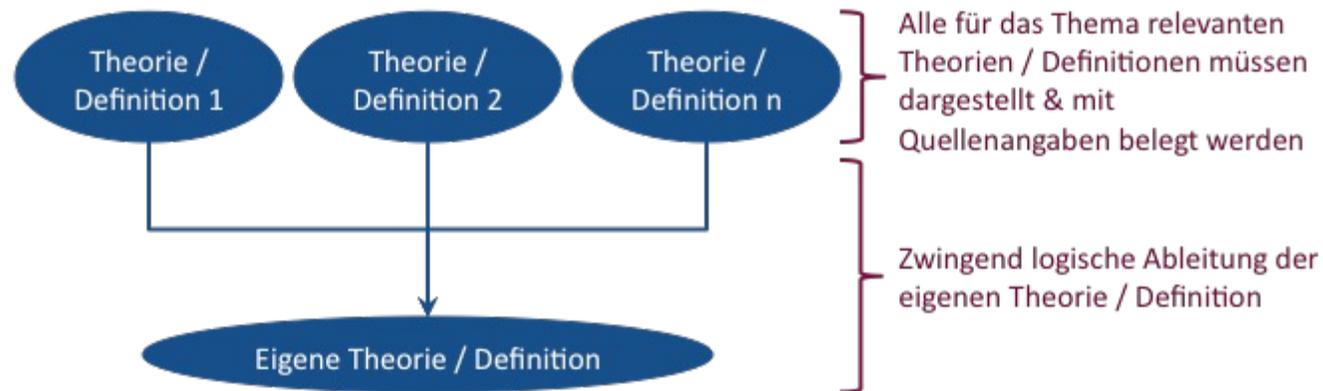
Methodik

Ergebnisse

Diskussion



- Theorie: Konzeptualisierung der zu untersuchenden Konstrukte



- Stand der Forschung: Antezedenzen, Konsequenzen & Korrelate der Konstrukte
 - Antezedenzen ~ Was sind die Ursachen für die Konstrukte?
 - Konsequenzen ~ Was sind die Folgen der Konstrukte?
 - Korrelate ~ Welche Zusammenhänge liegen für die Konstrukte vor?

- Gliederung
 - i.d.R. für jedes zu untersuchende Konstrukt einen Unterabschnitt & ggf. einen weiteren Unterabschnitt für den Untersuchungsbereich

- Beispiel: Determinanten von Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy
 - 1 Einleitung
 - 2 Theoretischer Hintergrund
 - 2.1 Kundenzufriedenheit und Kundenbindung
 - 2.2 Sharing Economy und Online-Plattformen
 - 2.3 Marktplatz-Akzeptanzmodell nach Kollmann
 - 2.4 Weitere Determinanten von Kundenzufriedenheit und Kundenbindung
 - 3 Hypothetisches Modell und Forschungsfragen
 - ...

Quelle: Godbersen, H., Brühl, R. (2019): Determinanten von Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy, in: transfer – Zeitschrift für Kommunikation und Markenmanagement, Jg. 65, Nr. 4, S. 6-17

- Zitationswürdige (wissenschaftliche) Literatur:
 - Lehr- und Handbücher (i.d.R. lediglich für grundlegende Definitionen geeignet)
 - Veröffentlichungen in Fachzeitschriften (i.d.R. zwingend erforderlich, insb. im Hinblick auf den Stand der Forschung; beachte: die Wissenschaftssprache ist Englisch)
- I.d.R. nicht zitationswürdige Literatur:
 - Populärwissenschaftliche Publikationen, „Beraterliteratur“, Internet, Lexika, Wörterbücher etc.
 - Faustregel: „Alles, was am Kiosk zu kaufen ist, ist nicht zitationswürdig.“
 - Ausnahmen (Internet, Zeitungsartikel etc.):
 - Ergebnisse von empirischen Erhebung mit nachvollziehbarer und belastbarer Methodik
 - Allgemeine Entwicklungen der Gesellschaft oder von Unternehmen

- (1) Suchen Sie nach potenziell relevanten Quellen für Ihr Forschungsprojekt auf Google Scholar
- (2) Prüfen Sie die Relevanz der gefundenen Quellen anhand der Abstracts
- (3) Beschaffen Sie sich die vollständigen Quellen
- (4) Arbeiten Sie die Quellen im Detail durch
- (5) Nutzen Sie die Extra-Funktionen von Google Scholar

The screenshot shows a Google Scholar search result for the article "The problem of abortion and the doctrine of double effect" by P. Foot (1967). The article is listed as an Oxford review. The abstract is partially visible. A red arrow points to the "Cited by 3040" and "Related articles" links. The interface includes a search bar, navigation tabs, and a sidebar with filters for "Any time", "Sort by", and "Any type".

Zusatzfunktionen in GoogleScholar:
„Cited by“ und „Related articles“

The screenshot shows a Google Scholar search result for "moral dilemma decision-making". The search results are sorted by relevance. The first result is "Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review" by J.E. Christensen and A. Gomila (2012). The abstract is partially visible. The interface includes a search bar, navigation tabs, and a sidebar with filters for "Any time", "Sort by", and "Any type".

Quelle: Godbersen, H., (2023): Qualitative Forschung für Dummies. Wiley.

Einleitung

Theorie & Stand der Forschung

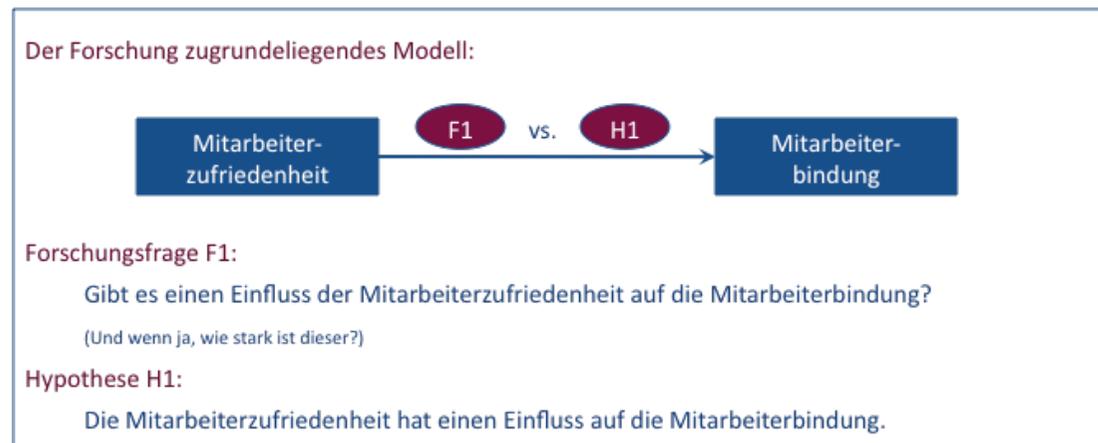
(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

- Ziel der quantitativen Forschung:
 - Durch die **Zuordnung von Zahlen zu empirischen Objekten oder Sachverhalten** (vgl. die „klassische“ Definition des Messens nach Orth) ein **vereinfachtes Abbild der Realität** (vgl. Definition von Modellen) zu geben und so **Theorien zu überprüfen**
- Forschungsfragen, Hypothesen & Modelle – vereinfachtes Beispiel:



- Die quantitative Forschung ist eine „Schwarz-Weiß-Welt“/“Ja-Nein-Welt“ → Dementsprechend müssen Forschungsfragen eindeutig beantwortbar sein / Hypothesen widerlegbar, bzw. falsifizierbar sein.

Hypothetisches Modell

- **Modell:** Ein Modell ist ein vereinfachtes Abbild der Realität.
- **Theorie:** Theorien bestehen aus Konstrukten (generalisierte Abstraktionen von Erscheinungen in der Realität) & deren Beziehungen zueinander.
- **Verhältnis von Modellen & Theorien:** Im wissenschaftlichen Kontext können Modelle und Theorien weitestgehend als Synonyme aufgefasst werden.
- **Empfehlung für die quantitative Forschung:** Wenn möglich, sollte ein (graphisches) Modell aus der Theorie & dem Stand der Forschung entwickelt werden & im empirischen Teil der Forschungsarbeit überprüft werden (“mit Zahlen gefüllt werden”) – Beispiel (Godbersen & Brühl, 2019):

Abb. 1: Theoretisches Modell – Determinanten der Kundenzufriedenheit und Kundenbindung

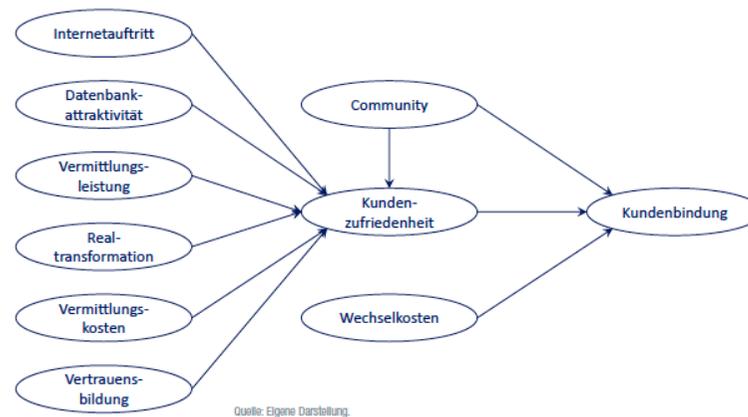
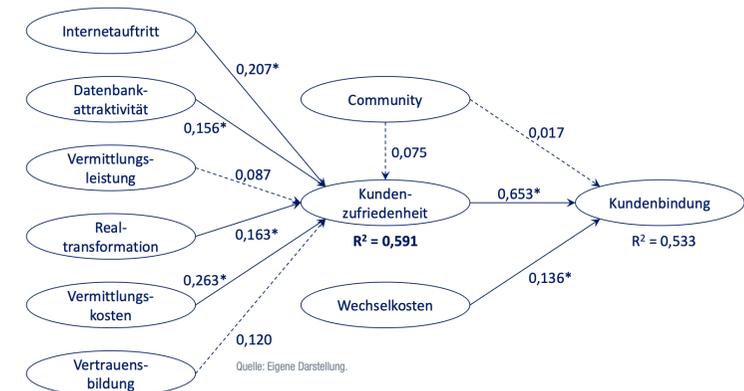
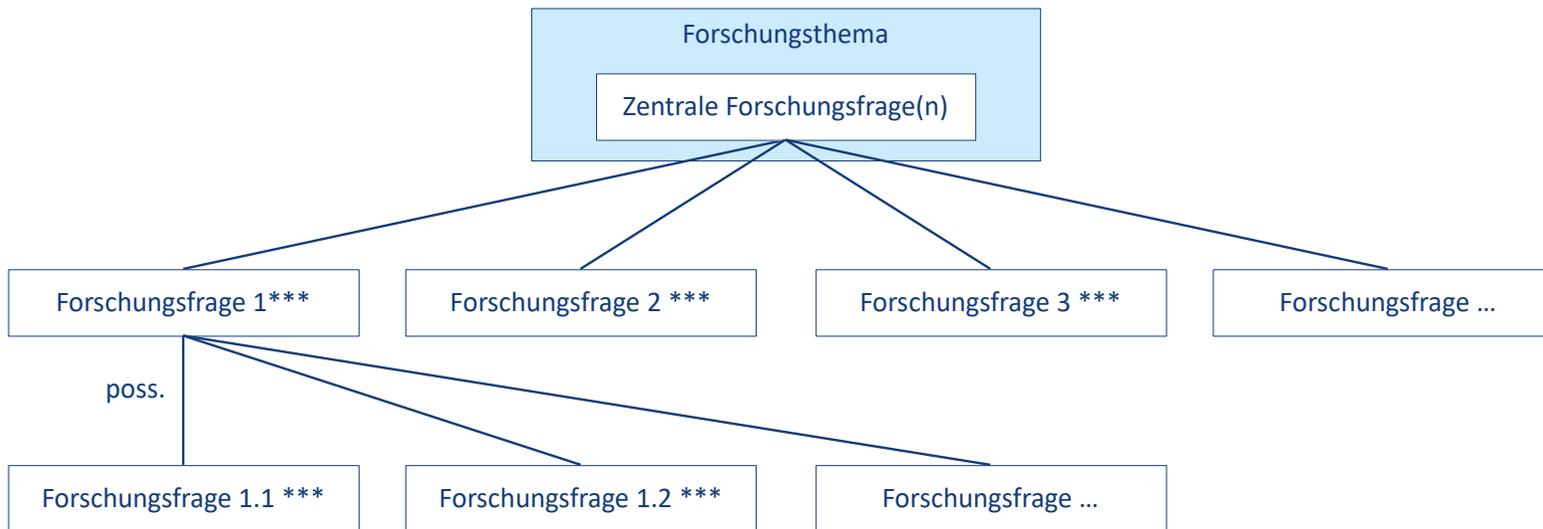


Abb. 3: Determinanten der Kundenzufriedenheit und Kundenbindung mit Pfadkoeffizienten und dem Bestimmtheitsmaß - Anm.: * p-Value < 0,05; n=246

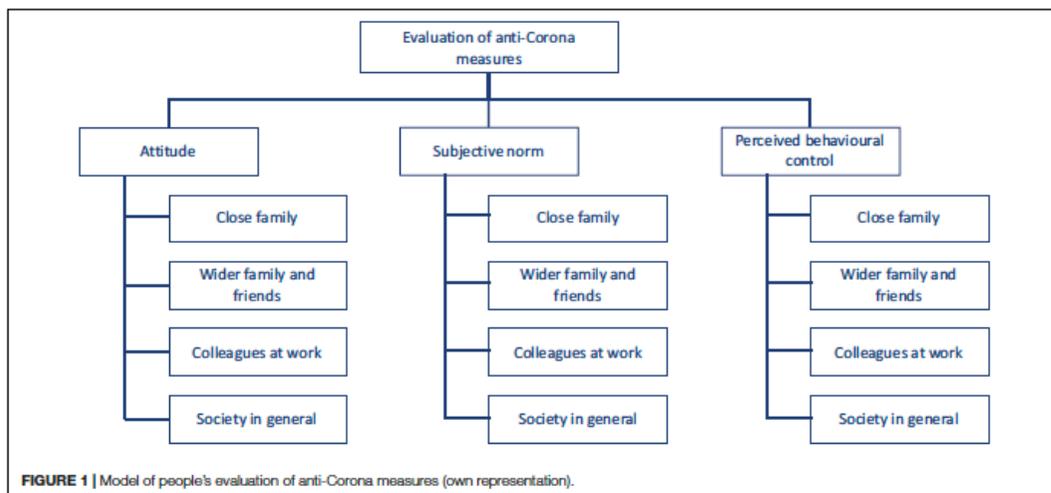




Empfehlung: Hierarchischer Aufbau der Forschungsfragen mit einer oder wenigen übergeordneten Forschungsfragen und mehreren untergeordneten Forschungsfragen (oder Hypothesen).

*** Anstelle von Forschungsfragen können in der quantitativen Forschung auch Hypothesen genutzt werden.

- Anforderungen an Forschungsfragen
 - Zwingend logisch und in erschöpfender Weise aus der Theorie und dem Stand der Forschung abgeleitet
 - Präzise und differenzierte formuliert
 - Forschungsfragen eindeutig beantwortbar



Source: Godbersen H, Hofmann LA and Ruiz-Fernández S (2020) How People Evaluate Anti-Corona Measures for Their Social Spheres: Attitude, Subjective Norm, and Perceived Behavioral Control. *Front. Psychol.* 11:567405. doi: 10.3389/fpsyg.2020.567405

RQ1 (subjective relevance): Which relevance do the attitude, subjective norm, and perceived behavioral control as well as the social spheres have for the people's evaluation of anti-Corona measures?

RQ1.1: Which relevance do the attitude, subjective norm, and perceived behavioral control have for the people's evaluation of anti-Corona measures?

RQ1.2: Which relevance do the social spheres—close family, wider family and friends, colleagues at work, and society in general—have for the people's evaluation of anti-Corona measures?

RQ2 (subjective quality): How well do people evaluate anti-Corona measures—restrictions on outdoor activities, tips for hygiene and tips for mental health—with regard to their attitude, subjective norm, and perceived behavioral control as well as their social spheres?

RQ2.1: How well do people evaluate anti-Corona measures—restrictions on outdoor activities, tips for hygiene, and tips for mental health—with regard to their attitude, subjective norm, and perceived behavioral control?

RQ2.2: How well do people evaluate anti-Corona measures—restrictions on outdoor activities, tips for hygiene, and tips for mental health—with regard to their social spheres?

RQ2.3: How well do people evaluate anti-Corona measures—restrictions on outdoor activities, tips for hygiene, and tips for mental health—overall?

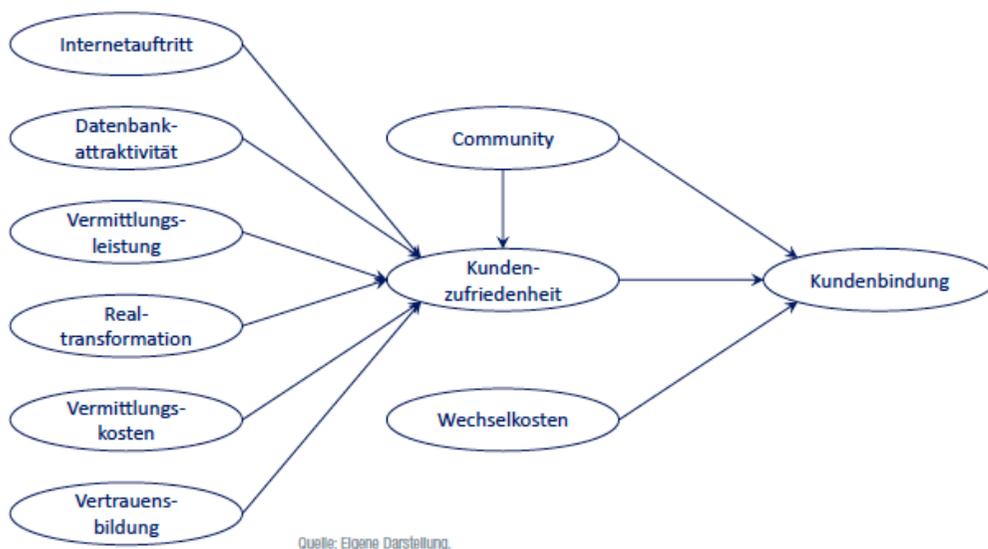
RQ3 (optimization): What is the potential of and the need for increasing the effectiveness of anti-Corona measures from the people's perspective, and with what priority should the current effectiveness of these measures be secured or increased with regard to attitude, subjective norm, and perceived behavioral control within the social spheres of people?

RQ3.1: From the people's perspective, what is the potential of and the need for increasing the effectiveness of anti-Corona measures with regard to attitude, subjective norm, and perceived behavioral control within the social spheres of people?

RQ3.2: Based on the potential of and need for increasing the effectiveness of anti-Corona measures, with which priority should the effectiveness of anti-Corona measures be secured or increased with regard to attitude, subjective norm, and perceived behavioral control within the social spheres of people?

Hypothetisches Modell & Forschungsfragen – Beispiele

Abb. 1: Theoretisches Modell – Determinanten der Kundenzufriedenheit und Kundenbindung



Quelle: Godbersen, H., Brühl, R. (2019): Determinanten von Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy, in: transfer – Zeitschrift für Kommunikation und Markenmanagement, Jg. 65, Nr. 4, S. 6-17

gangspunkt bildete Kollmanns (2001) Marktplatz-Akzeptanzmodell. Ausgehend davon werden der Internetauftritt, die Datenbankattraktivität, die Vermittlungsleistung, die Realtransformation und die Vermittlungskosten als Faktoren angenommen, die einen Einfluss auf die Kundenzufriedenheit ausüben (Abschnitt 2.3). Darüber hinaus wird davon ausgegangen, dass die Vertrauensbildung und die Community auf einer Online-Plattform weitere Determinanten der Kundenzufriedenheit darstellen (vgl. Abschnitt 2.4). Die Kundenbindung kann auf der Grundlage der im vorherigen Abschnitt behandelten Literatur als abhängig von der Kundenzufriedenheit (Abschnitt 2.1) sowie der Community und den Wechselkosten (Abschnitt 2.4) angesehen werden. Die soeben beschriebenen Zusammenhänge sind in >Abbildung 1 dargestellt.

Auf der Grundlage dieses Modells ergeben sich die folgenden Forschungsfragen:

F1: Welchen Einfluss üben der Internetauftritt, die Datenbankattraktivität, die Vermittlungsleistung, die Realtransformation, die Vermittlungskosten, die Vertrauensbildung

und die Community von Online-Plattformen der Sharing Economy auf die Kundenzufriedenheit aus?

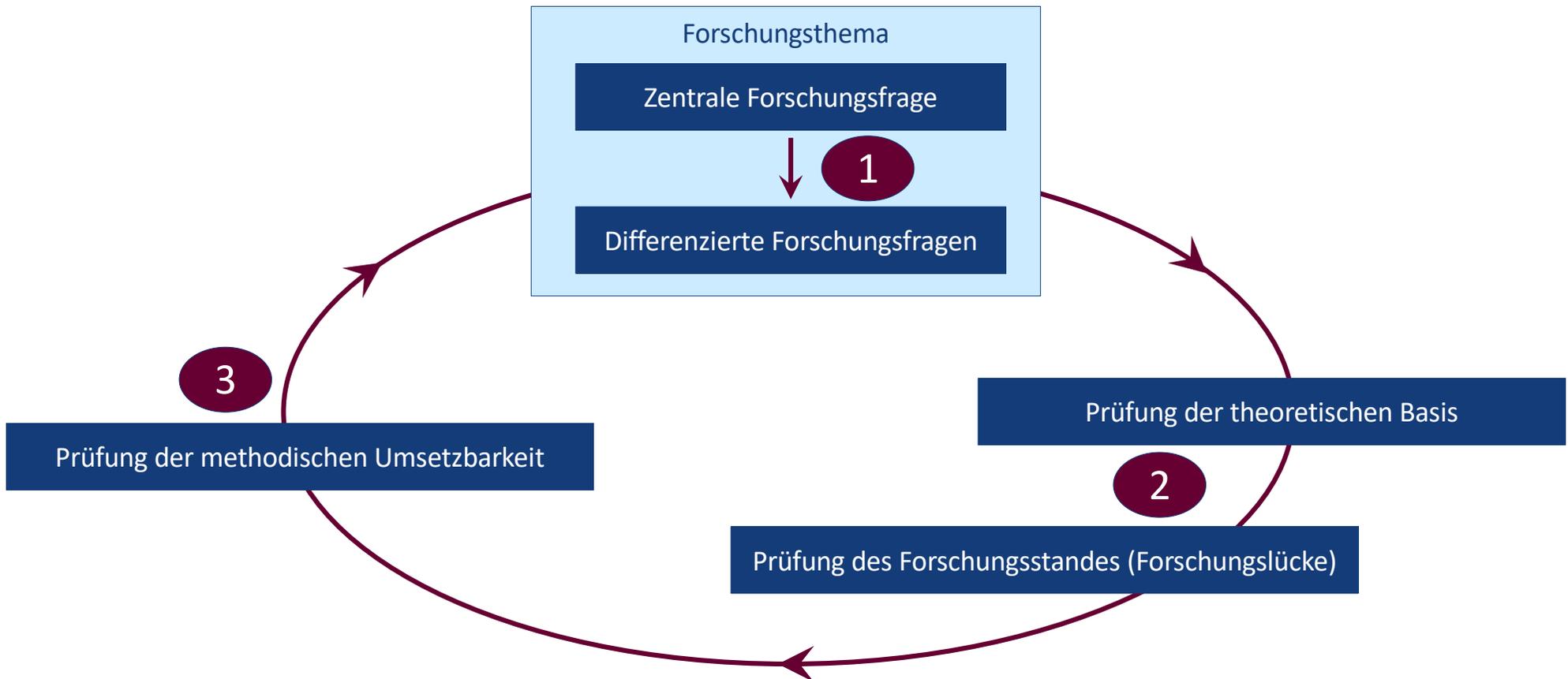
F2: Welchen Einfluss üben die Kundenzufriedenheit mit Plattformen der Sharing Economy, die Community auf entsprechenden Plattformen und die subjektiv wahrgenommenen Wechselkosten für die Nutzer auf die Kundenbindung aus?

4 Methode

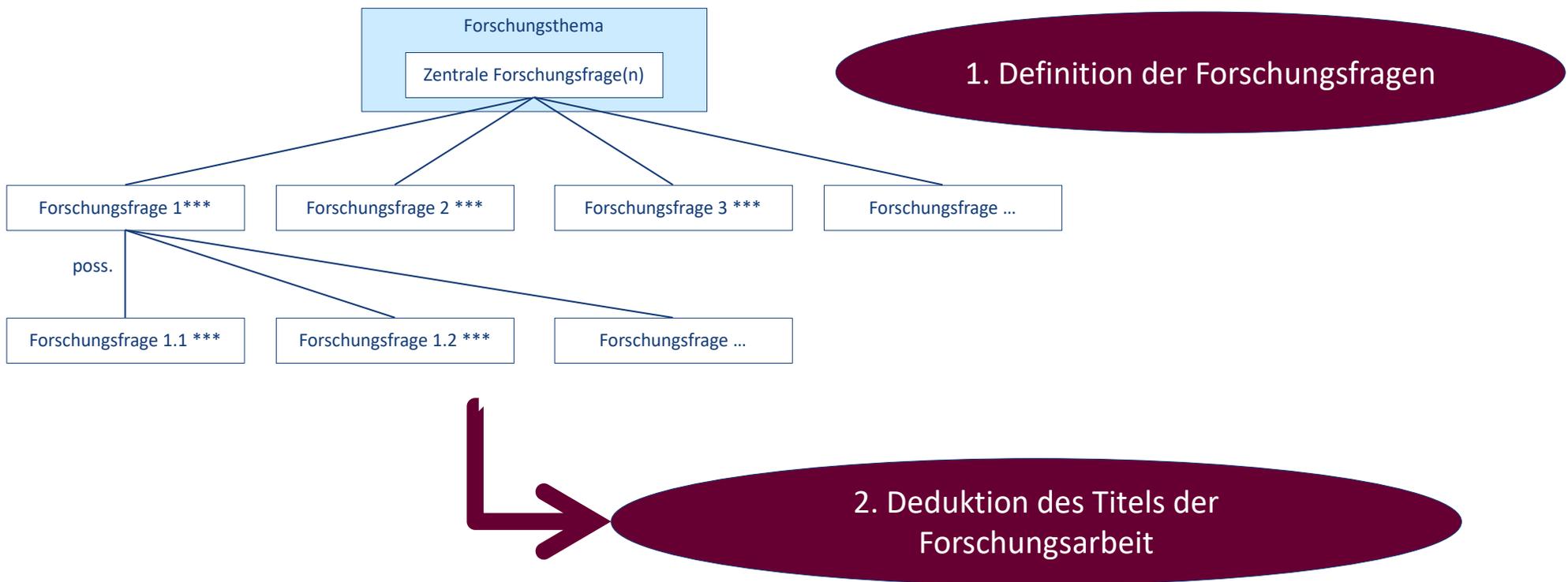
Im Folgenden werden das Untersuchungsdesign sowie die Messinstrumente und die Güteprüfung des Strukturmodells dargestellt.

4.1 Untersuchungsdesign

Um den Einfluss der in Abschnitt 3 dargestellten Faktoren auf die Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy zu untersuchen, wurde mit Airbnb ein konkreter Anbieter als Untersuchungsobjekt gewählt. Airbnb ermög-



Empfehlung: Definition der Forschungsfragen in einem zirkulären Prozess zu Beginn des Forschungsprojektes



Empfehlung: Erst die Forschungsfragen und dann den Titel als letzten Schritt der Vorbereitung auf die Forschungsarbeit entwickeln

Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Gliederung & Inhalte

Ergebnisse

Stichprobe & induktive Statistik

Diskussion

Skalenniveaus

Deskriptive Statistik

Inferenzstatistik

- Untersuchungsdesign (4.1)
 - Durchführung mit Erhebungszeitraum, Erhebungsmethode (Online, schriftlich, persönlich etc.), Rekrutierung der Probanden etc.
 - Stichprobe mit Umfang & Beschreibung nach soziodemographischen Variablen & Kontextvariablen
- Messinstrumente (4.2)
 - Fragebogen mit Inhalt (Items) & Skala
 - Auswertung mit statistischen Verfahren
 - Ggf. Güteprüfung der Messinstrumente
- Beachte: Bei Experimentalforschung kann die Gliederung abweichen
 - Participants / Materials (Stimuli & Apparatus) / Procedure & Design

Quelle: Godbersen, H., Brühl, R. (2019): Determinanten von Kundenzufriedenheit und Kundenbindung bei Online-Plattformen der Sharing Economy, in: transfer – Zeitschrift für Kommunikation und Markenmanagement, Jg. 65, Nr. 4, S. 6-17

4.2 Messinstrumente

Die Konstrukte des in Abschnitt 3 dargestellten Messmodells wurden durch Multi-Item-Messungen operationalisiert. Inhaltlich basieren die Items auf den Ausführungen in Abschnitt 2 und sind in [>Tabelle 1](#) dargestellt.

Die genannten Items wurden mit sechsstufigen Ratingskalen gemessen, deren Skalenniveaus wie folgt beschriftet waren: „stimme gar nicht zu“ (1), „stimme nicht zu“ (2), „stimme eher nicht zu“ (3), „stimme eher zu“ (4), „stimme zu“ (5) und „stimme voll zu“ (6). Die erhobenen Daten wurden mit einer Pfadmodellierung ausgewertet. Als Software kam SmartPLS in der Version 3.2.7 zum Einsatz (Ringle et al., 2015).

Die Güte der formativ gebildeten Konstrukte des Modells wurde im Hinblick auf die Multikollinearität getestet. Als Indikator diente der Variance Inflation Factor (VIF), der bei den entsprechenden Items Werte zwischen 1,204 und 2,023 aufweist ([>Tabelle 2](#)), was auf eine hohe Güte des Modells hinweist.

Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

Gliederung & Inhalte

Stichprobe & induktive Statistik

Skalenniveaus

Deskriptive Statistik

Inferenzstatistik

Stichprobe – Grundbegriffe

- Grundgesamtheit: Alle Personen, für welche die Ergebnisse der Studie gelten sollen
- Stichprobe (letztendlich ist die Netto-Stichprobe relevant)
 - Brutto-Stichprobe: Alle Personen, die zum Zweck der Datenerhebung kontaktiert werden
 - Netto-Stichprobe: Alle Personen, von denen Daten erfolgreich gewonnen wurden
- Vollerhebung: Alle Personen der Grundgesamtheit werden im Rahmen der Datenerhebung kontaktiert.
- Teilerhebung: Lediglich ausgewählte Personen der Grundgesamtheit werden im Rahmen der Datenerhebung kontaktiert.
- Repräsentative Stichprobe: Eine repräsentative Stichprobe liegt dann vor, wenn die Verteilung der forschungsrelevanten Merkmale in der Stichprobe der Verteilung der Merkmale in der Grundgesamtheit entspricht.



- Forschungsansätze
 - Eher explorativ ausgerichtete Designs, z.B. Untersuchung, ob grundsätzlich ein Effekt von A auf B besteht
 - keine repräsentative Stichprobe erforderlich
 - Faustregel: mind. 200 auswertbare Fragebögen
 - Eher repräsentativ ausgerichtete Designs, z.B. Sonntagsfrage („Welche Partei würden Sie wählen, wenn am nächsten Sonntag Bundestagswahl wäre?“)
 - repräsentative Stichprobe erforderlich
- Verfahren der Stichprobenziehung
 - Bewusste Stichprobenziehung
 - Convenience Sampling
 - Quotierte Stichprobenziehung
 - ...
 - Zufallsstichprobe (jedes Element der Grundgesamtheit hat die gleiche Wahrscheinlichkeit, in die Stichprobe zu gelangen)

- Arbeitsschritte
 - Festlegen der Quotierungsmerkmale
 - Bestimmen der Verteilung der Quotierungsmerkmale in der Grundgesamtheit
 - Festlegen des Stichprobenumfangs
 - Bestimmen des Anteils der Quotierungsmerkmals-Ausprägungen in der Stichprobe
- Beispiel

Age group	Population of England (2021)		Sample
	Abs.	Rel.	
20 to 29	7,102,321	23.88%	478
30 to 39	7,753,274	26.06%	521
40 to 49	7,166,966	24.09%	482
50 to 59	7,724,309	25.97%	519
Sum	29,746,870	100.00%	2,000

- Formeln

- Formel für unendliche Grundgesamtheit:
$$n = \frac{z^2 \times p \times (1 - p)}{\Delta p^2}$$

- Formel für endliche Grundgesamtheit:
$$n = \frac{z^2 \times p \times (1 - p)}{\Delta p^2 + \frac{z^2 \times p \times (1 - p)}{N}}$$

- Faustregel: Wenn $n/N < 0.05 \rightarrow$ Formel für unendliche Grundgesamtheit kann genutzt werden

- Variablen

- n = Stichprobenumfang
- N = Umfang der Grundgesamtheit
- z = z-value, abhängig vom Konfidenzniveau (mit welcher Wahrscheinlichkeit kann von der Stichprobe auf die Grundgesamtheit geschlossen werden; Standard: 95%)
- Δp = Stichprobenfehler (um wieviel Prozentpunkte weichen die gemessenen Werte von den wahren Werten in der Grundgesamtheit ab; Standardwert: 3%)
- p = Merkmalsverteilung in der Stichprobe (wenn unbekannt, 50%)

Stichprobenumfang bei unendlicher Grundgesamtheit

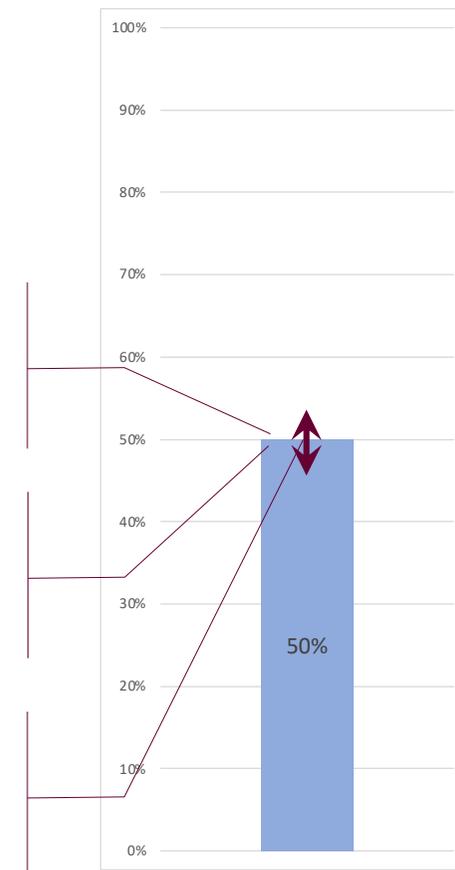
$$n = \frac{z^2 \times p \times (1 - p)}{\Delta p^2}$$

p (angen. Merkmalsverteilung)	50%	(1)
(1-p)	50%	
z (bei Konfidenzniveau 95%)	1.96	(2)
Δp (Stichprobenfehler)	3%	(3)
n (Stichprobenumfang)	1,067	(4)

(1) Annahme einer Merkmalsverteilung von 50%

(2) Mit einer Wahrscheinlichkeit von 95% stimmen die Werte der Stichprobe mit den wahren Werten der Grundgesamtheit überein

(3) Mit einer Wahrscheinlichkeit von 95% (2) liegen die wahren Werten der Grundgesamtheit im Bereich von +/- 3% des gemessenen Wertes (47% bis 53%)



- Stichprobenumfänge können mit Online-Tools berechnet werden

Quelle: <https://www.calculator.net/sample-size-calculator.html>.
Accessed: 2023-08-24

Sample Size Calculator

Modify the values and click the Calculate button to use

Find Out The Sample Size

This calculator computes the minimum number of necessary samples to meet the desired statistical constraints.

Confidence Level:	95%	
Margin of Error:	5	%
Population Proportion:	50	% Use 50% if not sure
Population Size:		Leave blank if unlimited population size.
<input type="button" value="Calculate"/>		<input type="button" value="Clear"/>

Find Out the Margin of Error

This calculator gives out the margin of error or confidence interval of observation or survey.

Confidence Level:	95%	
Sample Size:	100	
Population Proportion:	60	%
Population Size:		Leave blank if unlimited population size.
<input type="button" value="Calculate"/>		<input type="button" value="Clear"/>

- Nach der Datenerhebung kann die Formel für die Berechnung des Stichprobenumfangs umgestellt werden, so dass bei einer gegebenen Merkmalsverteilung das Konfidenzniveau & der Stichprobenfehler neu bestimmt werden können.
- Induktive Statistik –Kernfrage:
 - Mit welcher Wahrscheinlichkeit und welchem Stichprobenfehler („Trefferbereich“) kann von der Stichprobe auf die Grundgesamtheit geschlossen werden?

$$n = \frac{z^2 \times p \times (1 - p)}{\Delta p^2}$$
$$n = \frac{z^2 \times p \times (1 - p)}{\Delta p^2 + \frac{z^2 \times p \times (1 - p)}{N}}$$

Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

Gliederung & Inhalte

Stichprobe & induktive Statistik

Skalenniveaus

Deskriptive Statistik

Inferenzstatistik

Skalenniveau		Math. Eigenschaften	Beschreibung der Messwertcharakteristika	Beispiel	Lageparameter
Kategorial	Nominal	\neq	Die Messwerte sind identisch oder nicht identisch.	Geschlecht	<ul style="list-style-type: none"> • Modus
	Ordinal	\neq ; $</>$	Die Messwerte sind größer, kleiner oder identisch.	Olympische Ränge	<ul style="list-style-type: none"> • Modus • Median
Metrisch	Intervall	\neq ; $</>$; - ; +	Die Distanz zwischen den Messwerten kann angegeben werden.	Temperatur	<ul style="list-style-type: none"> • Modus • Median • Arithm. Mittel
	Verhältnis	\neq ; $</>$; + / - ; * / ÷	Die Distanz zwischen den Messwerten und deren Verhältnis kann angegeben werden.	Körpergröße	

- Nominales Skalenniveau

Welches Geschlecht haben Sie?

Männlich

Weiblich

Divers

- Ordinales Skalenniveau

Bitte bewerten Sie die folgenden Produkte, indem Sie diese in eine Reihenfolge von 1 (bestes Produkt) bis 3 (schlechtestes Produkt) bringen.

Produkt A _____

Produkt B _____

Produkt C _____

- Metrisches Skalenniveau

Wie alt sind Sie in Jahren? _____

- Ratingskala (streng genommen ordinales Skalenniveau; kann aber metrisch behandelt werden)

Wie zufrieden sind Sie mit Produkt XY auf einer Skala von 1 „überhaupt nicht zufrieden“ bis 6 „voll und ganz zufrieden“?

überhaupt nicht zufrieden voll und ganz zufrieden

Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

Gliederung & Inhalte

Stichprobe & induktive Statistik

Skalenniveaus

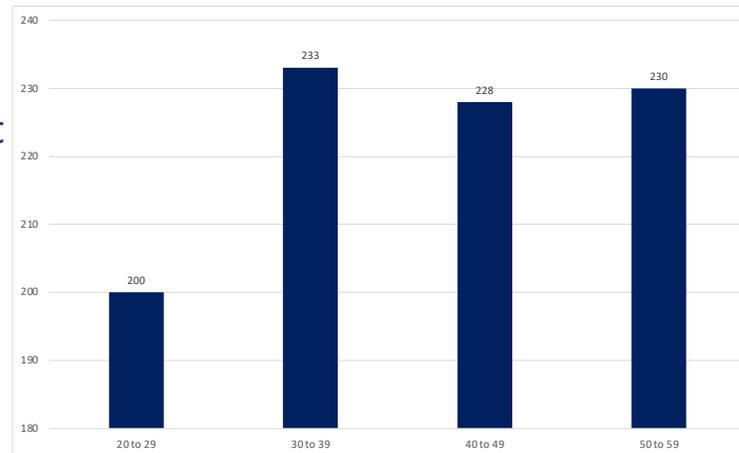
Deskriptive Statistik

Inferenzstatistik

- Die deskriptive Statistik zielt darauf ab, eine Stichprobe / Datenmenge zu beschreiben.
- Wesentliche statistische Verfahren:
 - Häufigkeiten
 - Lageparameter
 - Streuungsmaße
- Beachte:
 - In der wissenschaftlichen Forschung dient die deskriptive Statistik vornehmlich dazu, die Stichprobe nach soziodemographischen Variablen und Kontextvariablen zu beschreiben. Eine Forschungsarbeit, die ausschließlich auf deskriptiven Analysen basiert, wird wissenschaftlichen Anforderungen i.d.R. nicht gerecht.

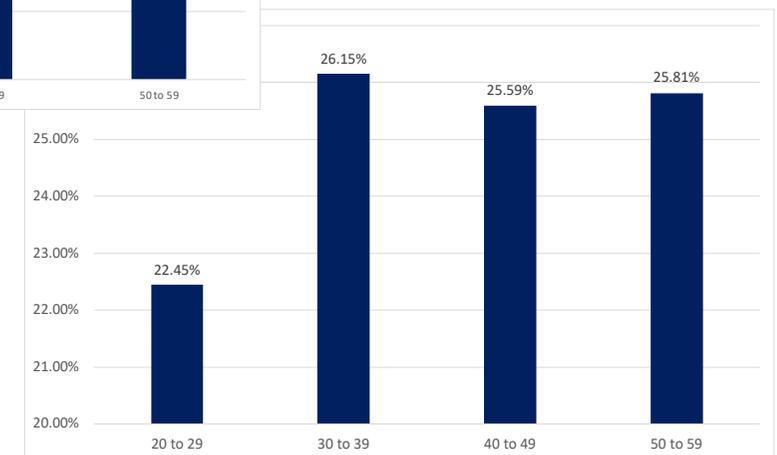
Häufigkeiten: Steckbrief

- Absolute Häufigkeit:
 - Gibt an wie häufig ein Wert auftritt



- Relative Häufigkeit:
 - Gibt das Verhältnis eines untersuchten Merkmalsträgers zu einem Wert oder einer Klasse von Werten an

Age group	female	male	total
20 to 29	22.08%	22.84%	22.45%
30 to 39	28.35%	23.78%	26.15%
40 to 49	23.16%	28.21%	25.59%
50 to 59	26.41%	25.17%	25.81%
SUM	100.00%	100.00%	100.00%



- Kreuztabelle:
 - Darstellung der gemeinsamen Verteilung zweier Variablen

- Arithmetisches Mittel:
 - Division der Summe aller Werte durch die Anzahl der Werte; metrisches Skalenniveau notwendig
- Median:
 - Wert, der die Grenze zwischen unterer und oberer Hälfte markiert; mindestens ordinales Skalenniveau
- Modalwert/Modus:
 - Häufigste Wert einer Häufigkeitsverteilung; mindestens nominales Skalenniveau

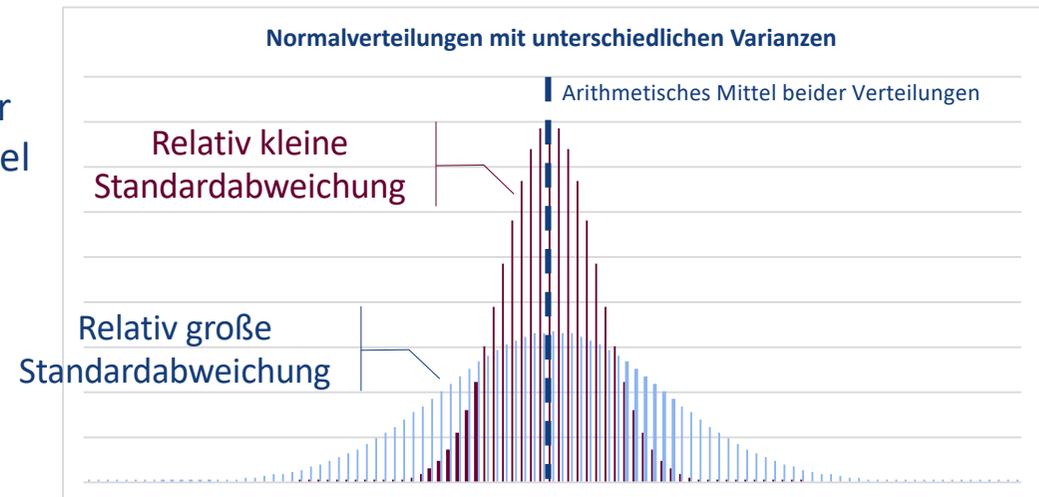
Proband	Variable A
1	23
2	27
3	30
4	35
5	40
6	55
7	60
8	60
9	75
Summe	405
Arith. Mittel	45

Median

Modus

- Varianz
 - Durchschnittliche quadrierte Abweichung der gemessenen Werte vom arithmetischen Mittel (beachte: empirische Varianz \rightarrow /n ; Stichprobenvarianz $\rightarrow /n-1$)

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$



- Standardabweichung
 - Durchschnittliche Abweichung der gemessenen Werte vom arithmetischen Mittel

$$\sigma = \sqrt{\sigma^2}$$

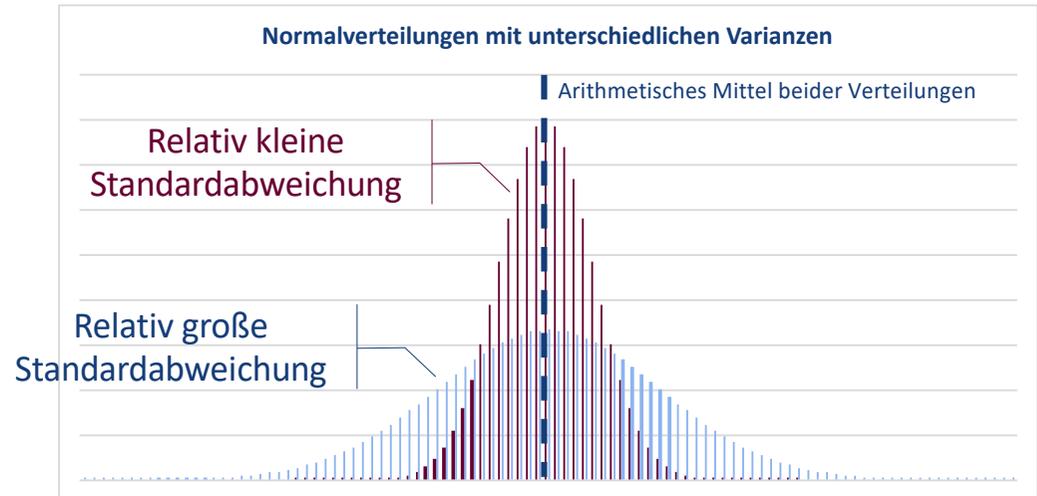
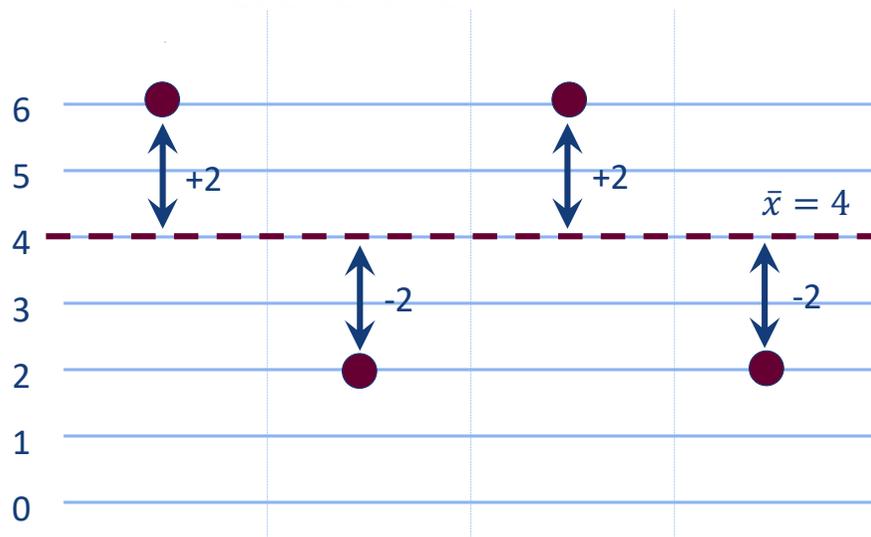
- **Streuungsparameter**

- Varianz

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

- Standardabweichung

$$\sigma = \sqrt{\sigma^2}$$



$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x}) &= \{ (+2) + (-2) + (+2) + (-2) = 0 \} \\ \sum_{i=1}^n (x_i - \bar{x})^2 &= \{ (+2)^2 + (-2)^2 + (+2)^2 + (-2)^2 = 4 + 4 + 4 + 4 = 16 \} \\ \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} &= \{ 16 / 4 = 4 \} \\ \sigma = \sqrt{\sigma^2} &= \{ \sqrt{4} = 2 \} \end{aligned}$$

Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

Gliederung & Inhalte

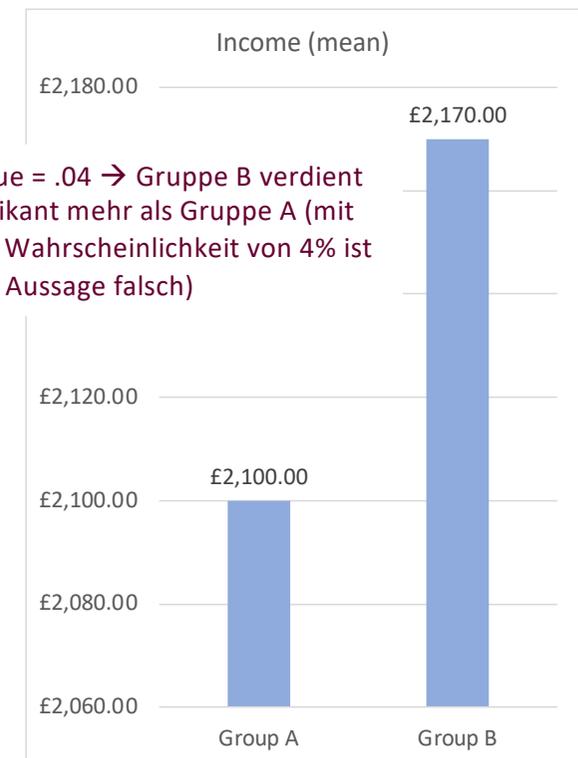
Stichprobe & induktive Statistik

Skalenniveaus

Deskriptive Statistik

Inferenzstatistik

- Mit Inferenzstatistik wird bestimmt, ob gemessene Effekte signifikant sind. Dabei bedeuten:
 - Signifikant = nicht-zufällig (umgangssprachlich vereinfacht: es gibt „wirklich“ einen Effekt)
 - Nicht-signifikant = zufällig (umgangssprachlich vereinfacht: es gibt keinen Effekt)
- Die quantitative „Schwarz-Weiß-Welt“
 - H1: Effekt XY liegt vor.
 - H0: Effekt XY liegt nicht vor.
- „Logik“ der inferenzstatistischen Verfahren
 - Die inferenzstatistischen Verfahren versuchen H0 zu verwerfen.
 - Das Ergebnis ist der p-value mit Werten zwischen 0 und 1.
 - Der p-value sagt aus, mit welcher Wahrscheinlichkeit H0 gilt, obwohl H0 verworfen wurde.
- Vorgehen bei der Analyse
 - (1) p-value → Signifikanz: $p < .05$ = signifikant, $p < .01$ = hochsignifikant, $p < .001$ = höchstsignifikant
 - (2) Bei p-value $< .05$: gemessene Werte (arith. Mittel, Häufigkeiten etc.) anschauen → Richtung & Stärke des Effekts



Skalenniveau		Math. Eigenschaften	Beschreibung der Messwertcharakteristika	Beispiel	Lageparameter
Kategorial	Nominal	\neq	Die Messwerte sind identisch oder nicht identisch.	Geschlecht	<ul style="list-style-type: none"> • Modus
	Ordinal	\neq ; $</>$	Die Messwerte sind größer, kleiner oder identisch.	Olympische Ränge	<ul style="list-style-type: none"> • Modus • Median
Metrisch	Intervall	\neq ; $</>$; - ; +	Die Distanz zwischen den Messwerten kann angegeben werden.	Temperatur	<ul style="list-style-type: none"> • Modus • Median • Arithm. Mittel
	Verhältnis	\neq ; $</>$; + / - ; * / ÷	Die Distanz zwischen den Messwerten und deren Verhältnis kann angegeben werden.	Körpergröße	

Strukturprüfende Verfahren		Unabhängige Variable	
		nicht-metrisch	metrisch
Abhängige Variable	nicht-metrisch	Chi ² -Test	(Diskriminanzanalyse)
	metrisch	t-Test*** (bei 2 Gruppen) & Varianzanalyse (bei ≥ 3 Gruppen)	Regressionsanalyse (Dependenz) & Korrelationsanalyse (Interdependenz)

*** „Ergänzende“ Tests:

- Shapiro-Wilk-Test
(prüft Stichproben auf Normalverteilung, die für den t-Test notwendig ist)
- Mann-Whitney-Test/Wilcoxon-Test
(prüft mind. ordinal skalierte Daten auf Unterschiede ohne Berücksichtigung der Verteilungsform)

Strukturentdeckende Verfahren	<ul style="list-style-type: none"> • Hauptkomponentenanalyse (Faktoranalyse) • (Clusteranalyse)
-------------------------------	---

- **Anwendung:**
 - Testen auf (Un-)Abhängigkeit bei nominalen Daten (Weichen die gemessenen Werte nicht-zufällig von den erwarteten Werten ab?)
- **Fragestellung:**
 - Gibt es einen „nicht-zufälligen“ statistischen Zusammenhang zwischen zwei nominalskalierten Merkmalen? Sind zwei kategoriale Variablen unabhängig?
- **Vorgehen bei der Analyse:**
 - 1) Ist der p-value <0.05 (<0.01 , <0.001)? → Signifikanz?
 - 2) Wie unterscheiden sich die Häufigkeiten?
- **„Statistische“ Analyseschritte („im Hintergrund“):**
 - 1) Berechnung der erwarteten Häufigkeiten (e), wenn kein Unterschied zwischen den Variablen bestehen würde (Spaltensumme * Zeilensumme / Merkmalsträger)
 - 2) Bestimmung, ob die gemessenen Werte (h) signifikant von den erwarteten Werten abweichen?

	männlich	weiblich	Summe
blau	h = 35 e = 25	h = 15 e = 25	50
rot	h = 15 e = 25	h = 35 e = 25	50
Summe	50	50	100

- „Statistische“ Analyseschritte („im Hintergrund“):

- 1) Berechnung der erwarteten Häufigkeiten (e), wenn kein Unterschied zwischen den Variablen bestehen würde (Spaltensumme * Zeilensumme / Merkmalsträger)
- 2) Bestimmung, ob die gemessenen Werte (h) signifikant von den erwarteten Werten abweichen?

V1	männlich	weiblich	Summe
blau	h = 15 e = 25	h = 35 e = 25	50
rot	h = 35 e = 25	h = 15 e = 25	50
Summe	50	50	100

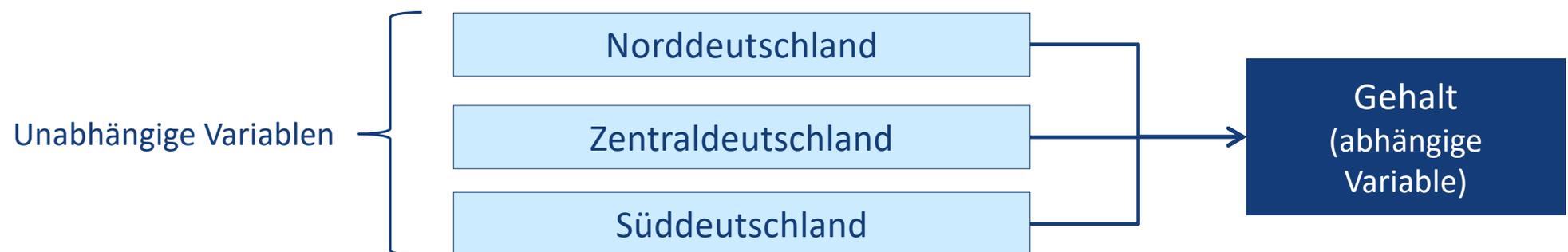
p-value = 0.0001447

V2	männlich	weiblich	Summe
blau	h = 24 e = 25	h = 26 e = 25	50
rot	h = 26 e = 25	h = 24 e = 25	50
Summe	50	50	100

p-value = 0.8415

- **Anwendung:**
 - Der t-Test ist ein parametrisches Verfahren. Beim Zweistichproben-t-Test wird überprüft, ob sich eine metrische Variable zwischen zwei Stichproben signifikant (nicht zufällig) unterscheidet.
- **Fragestellung:**
 - Gibt es einen „nicht-zufälligen“ statistischen Unterschied zwischen zwei Stichproben in Bezug auf ein metrisches Merkmal?
 - Gibt es einen „nicht-zufälligen“ Einfluss einer nominalen Variable (2 Gruppen) auf eine metrische Variable?
- **Differenzierung zwischen:**
 - t-Test bei unabhängigen Stichproben – Messungen bei zwei Gruppen
 - t-Test bei abhängigen Stichproben – Messungen bei derselben Gruppe (z.B. vorher – nachher)
 - Einstichproben-t-Test (Vergleich eines Mittelwertes mit einem vorgegebenen Wert)
- **Vorgehen bei der Analyse:**
 - 1) Ist der p-value <0.05 (<0.01 , <0.001)? → Signifikanz?
 - 2) Wie unterscheiden sich die Mittelwerte?

- **Verwendung:**
 - Die Varianzanalyse (*Analysis of Variance, ANOVA*) ist ein parametrisches Verfahren und untersucht, ob sich die Mittelwerte mehrerer (mehr als zwei) unabhängiger Stichproben systematisch unterscheiden.
- **Fragestellung:**
 - Gibt es einen „nicht-zufälligen“ statistischen Unterschied zwischen mehr als zwei Stichproben in Bezug auf ein metrisches Merkmal?
- **Abgrenzung zum t-Test:**
 - Während der t-Test metrische Merkmale von maximal zwei Stichproben auf signifikante Unterschiede prüft, werden bei der Varianzanalyse mehrere Mittelwerte verglichen
- **Beispiel:**



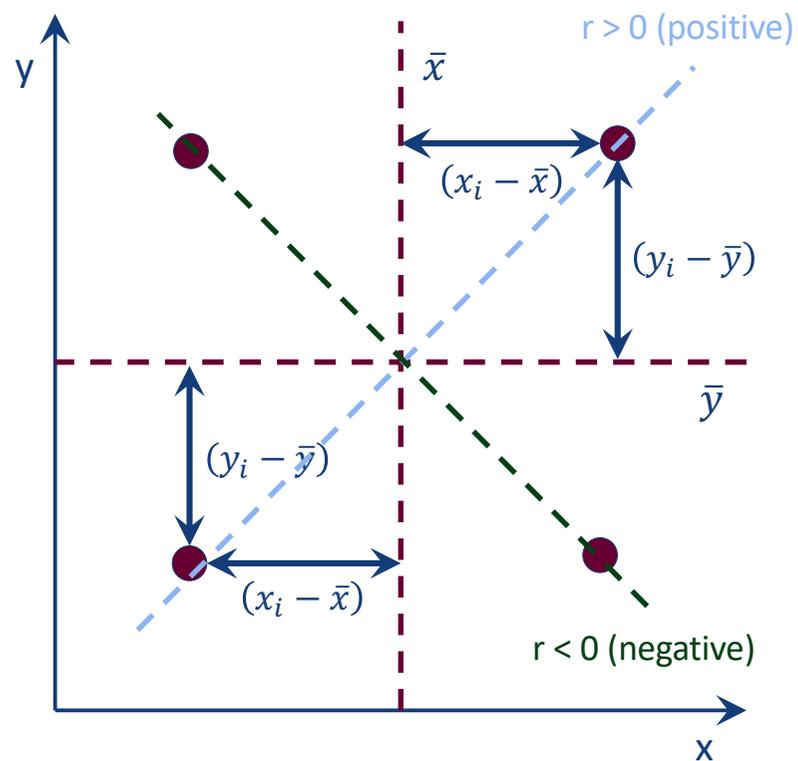
- **Verwendung:**
 - Die Korrelationsanalyse untersucht den (linearen) Zusammenhang zwischen zwei oder mehr metrischen Variablen.
- **Fragestellung:**
 - Wie stark ist der lineare Zusammenhang zwischen den Variablen – und in welcher Richtung besteht er?
→ Wird der Wert von Variable A erhöht / gesenkt oder bleibt er gleich, wenn sich der Wert von Variable B ändert?
- **Vorgehen:**
 - 1) Ist der p-value < 0.05 (< 0.01 , < 0.001)? → Signifikanz?
 - 2) Wie hoch ist der Korrelationskoeffizient?
- **Mögliche Interpretation des Korrelationskoeffizienten:**

$0 < r \leq +1.0$	gleichgerichteter Zusammenhang
$-1.0 \geq r < 0$	entgegengesetzter Zusammenhang

$0.2 < r \leq 0.5$	geringe Korrelation
$0.5 < r \leq 0.7$	mittlere Korrelation
$0.7 < r \leq 1.0$	hohe Korrelation

Korrelation: Steckbrief

- Statistische Erklärung des Korrelationskoeffizienten



Kovarianz (gemeinsame Streuung zweier Variablen):

$$\text{cov} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

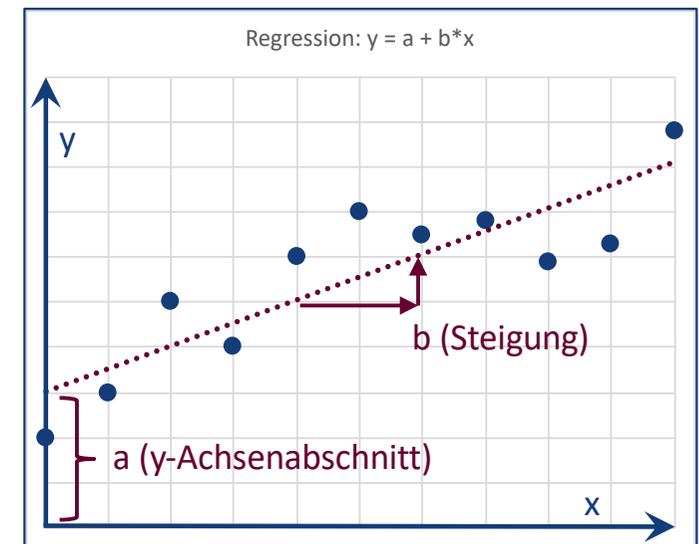
Varianz (Streuung einer Variablen):

$$\text{var} = \sigma_x = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Korrelationskoeffizient (Anteil der gemeinsamen Streuung zweier Variablen an der gesamten Streuung):

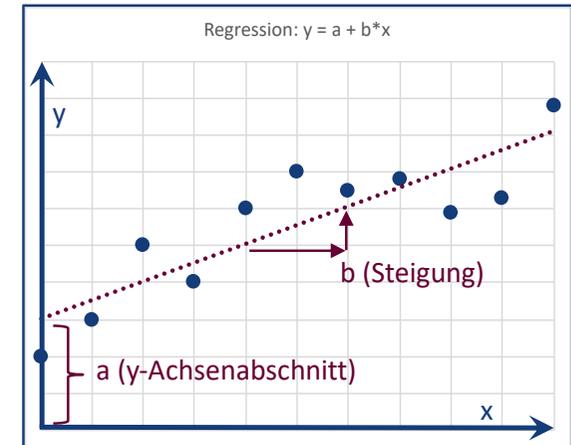
$$r = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\text{var}_x * \text{var}_y}}$$

- **Verwendung:**
 - Die lineare Regression untersucht den linearen Einfluss einer metrischen Variable (Einfachregression) oder mehrerer metrischer Variablen (Multiple Regression) auf eine abhängige metrische Variable.
- **Fragestellung:**
 - Inwieweit kann der Wert einer abhängigen Variablen durch die Werte einer (Einfachregression) oder mehrerer (Multiple Regression) unabhängiger Variablen erklärt werden?
- **Regressionsgerade (Ergebnis der Analyse):**
 - $y = a + b \cdot x$
 - y: Abhängige Variable („Wirkung“)
 - x: Unabhängige Variable („Ursache“)
 - a: y-Achsenabschnitt
 - b: Regressionskoeffizient („um wieviel steigt y, wenn sich x um eine Einheit erhöht“)
- **Statistische Analyse („im Hintergrund“):**
 - Im Rahmen der Regressionsanalyse wird die Gerade ermittelt, bei der die quadrierten Abweichungen der gemessenen Werte von der Geraden am geringsten ist (Methode der kleinsten Quadrate)



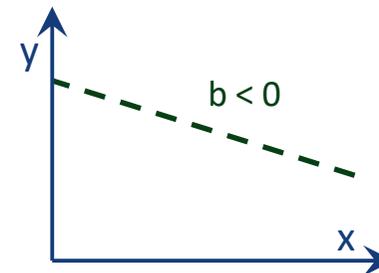
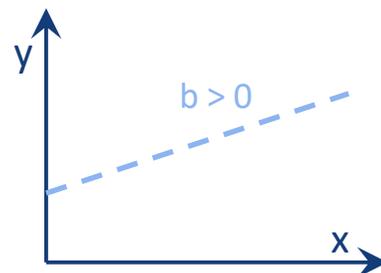
- **Vorgehen:**

- 1) Sind die p-values < 0.05 (< 0.01 , < 0.001)?
 - Liegt ein signifikanter Einfluss von x_i auf y vor (p-values)?
- 2) Wie hoch sind die Regressionskoeffizienten (b)?
 - Wie stark ist der Einfluss von x_i auf y ?
- 3) Welchen Wert hat das Bestimmtheitsmaß (R^2)?
 - Wie hoch ist der Erklärungsgehalt des Modells?
 - R^2 = Anteil der Varianz der abhängigen Variable, den die unabhängige(n) Variable(n) erklären



- **Beachten Sie:**

- Der Regressionskoeffizient kann positive und negative Werte annehmen (vgl. auch Korrelation)



Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

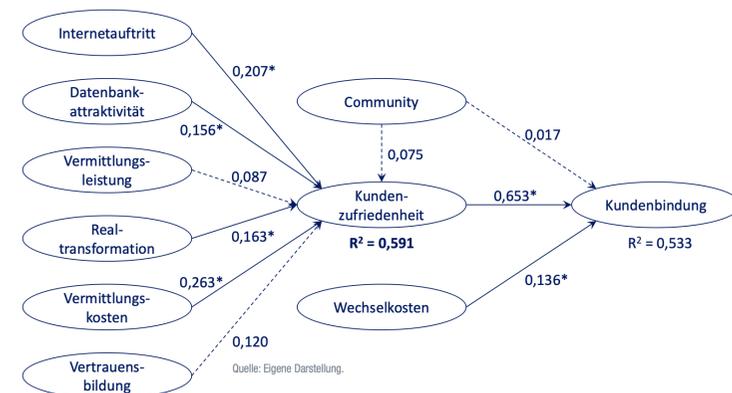
- „Neutrales & objektives“ Berichten der Ergebnisse
- I.d.R. deskriptive Statistik der untersuchten Konstrukte
- I.d.R. Ergebnisse der Inferenzstatistik mit Effektgrößen, p-values & ggf. weiteren Indikatoren (ohne weiterführende Interpretationen)

Tab. 5: Deskriptive Statistik der gemessenen Konstrukte (n=246)

Konstrukte	Min.	Max.	Mittelwert	Standardabweichung
Internetauftritt	2,50	6,00	4,79	0,69
Datenbankattraktivität	2,75	6,00	4,82	0,61
Vermittlungsleistung	2,00	6,00	4,84	0,70
Realtransformation	2,00	6,00	4,78	0,70
Vermittlungskosten	1,00	6,00	4,09	1,06
Vertrauensbildung	1,75	6,00	4,87	0,71
Community	1,00	6,00	3,21	0,98
Wechselkosten	1,00	6,00	3,55	1,05
Kundenzufriedenheit	2,00	6,00	4,90	0,73
Kundenbindung	2,00	6,00	5,36	0,77

Quelle: Eigene Darstellung.

Abb. 3: Determinanten der Kundenzufriedenheit und Kundenbindung mit Pfadkoeffizienten und dem Bestimmtheitsmaß - Anm.: * p-Value < 0,05; n=246



Einleitung

Theorie & Stand der Forschung

(Hypothetisches Modell &) Forschungsfragen

Methodik

Ergebnisse

Diskussion

- Die Diskussion besteht i.d.R. aus zwei Unterabschnitten
 - Implikationen (6.1)
 - Limitationen & Ausblick (6.2)
- Leitfragen beim Schreiben der Diskussion
 - Implikationen (~Interpretation)
 - Welche inhaltlichen Erklärungen liegen hinter den Ergebnissen („Zahlen“)?
 - Was bedeuten die Ergebnisse für die Wissenschaft & die Praxis?
 - Wie verhalten sich die Ergebnisse zu Befunden aus anderer Forschung?
 - Limitationen & Ausblick
 - Welche Grenzen weist die vorliegende Forschung auf?
 - Welche Forschungsprojekte sollten sich dem vorliegenden Projekt anschließen?

Prof. Dr. Hendrik Godbersen

www.godbersen.online

Two Final Golden Rules for (Research) Life

Prof. Dr.
Godbersen

Get the basics right...

...and do the little things right.

and

Discipline & determination

or

Courage, determination, unselfishness & joyfulness in the face of adversity

