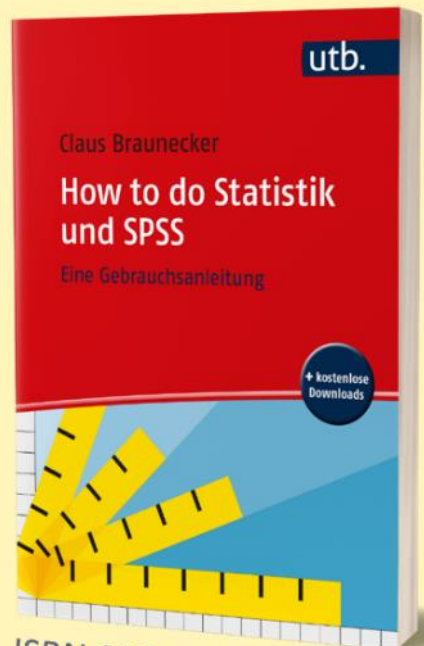


**Empirie –
leicht wie nie.**



ISBN 978-3-8252-5595-4

**Statistik und SPSS –
ganz ohne Stress.**



ISBN 978-3-8252-5596-1

9 | Umsetzungs-Tipps für die Praxis

... in diesem Kapitel geht's um:

● Umsetzungs-Tipps für Leitfaden- oder qualitative Interviews (9.1, S. 1)
● Umsetzungs-Tipps für qualitative Gruppendiskussionen (9.2, S. 2)
● Umsetzungs-Tipps für die Sekundäranalyse von Firmendaten (9.3, S. 3)
● Umsetzungs-Tipps für Inhaltsanalysen (9.4, S. 5 9.5, S. 6 9.6, S. 7)
● Umsetzungs-Tipps für verdeckte Testkäufe (Mystery-Shoppings) (9.7, S. 9)
● Umsetzungs-Tipps für „ Vor-Ort-Erhebungen “ (9.8, S. 10)
● Umsetzungs-Tipps für Mitarbeiterbefragungen (9.9, S. 12)
● Umsetzungs-Tipps für eine Potenzialanalyse (9.10, S. 14)
● Umsetzungs-Tipps für Expertenbefragungen (9.11, S. 16)
● Beispiel einer kaum erreichbaren Zielgruppe (9.12, S. 18)
● Umsetzungs-Tipps für ein Laborexperiment (9.13, S. 19)
● Umsetzungs-Tipps für ein experimentelles Befragungsdesign (9.14, S. 20)

Die folgenden Best-Practise-Beispiele orientieren sich an sehr vielen Ausführungen im Buch BRAUNECKER, CLAUS (2021): HOW TO DO EMPIRISCHE SOZIALFORSCHUNG. EINE GEBRAUCHSANLEITUNG. WIEN: FACULTAS.¹ Sie thematisieren unterschiedlichste Erkenntnisinteressen und methodische Settings (inkl. möglicher Unstimmigkeiten im Forschungsprozess).²

9.1 | Entwicklung eines idealen Fachbuchs

Ein Autorenkollegium möchte ein ideales Fachbuch für eine bestimmte studentische Zielgruppe entwickeln.

Die Methode erster Wahl stellen in diesem Fall qualitative Einzelgespräche dar. Sie erlauben es, die individuellen Aspekte, die hier zum Tragen kommen, detailliert zu erfassen. Da jemand im Regelfall alleine liest, ist die Erhebung bei Einzelpersonen gegenüber der ansonsten ebenso denkbaren Methodik von Fokusgruppen der idealere Weg. Je nachdem, wie „offen“ die Erhebung gestaltet werden soll, gelangen Gesprächsleitfaden (Leitfadeninterview) bis hin zu qualitativem Fragebogen zum Einsatz.

Zunächst muss die Zielgruppe, die das Buch anspricht, genau definiert werden. Die Grundgesamtheit wird soziodemografisch und auch nach ihrer „Fachbuchverwendung“ beschrieben. Das könnte z.B. über die Studienrichtung (Sozial-, Wirtschafts-, Naturwissenschaften ...), Ausbildungsstätte (Universität oder FH), Zeitaufwand (Fulltime- oder berufs begleitendes Studium) oder bereits absolvierte Semesteranzahl (Bakkalaureat- oder Master-Studierende) erfolgen.

¹ Alle Seitenverweise in eckigen Klammern beziehen sich auf dieses Werk.

² Datenschutzrechtliche Einschränkungen [21 ff.] finden dabei keine Beachtung.

Eine völlig andere – alternative oder ergänzende – Zielgruppe könnten Marketingfachleute („Expertinnen und Experten“) in Sach- und Fachbuchverlagen darstellen.

Da bei qualitativen Designs ein repräsentatives Sample meist wenig Sinn macht, hat diese Vorsichtung den primären Zweck, auf keine wichtige Teilzielgruppe zu vergessen.

Die Stichprobe braucht hier nicht zufällig zu sein – vielmehr bieten sich Quoten nach den oben definierten Segmenten an. Im Rahmen der qualitativen Einzelgespräche ist es damit möglich, ein detailliertes Bild eines idealen Fachbuchs zu zeichnen. Gibt es sehr spezielle oder zielgruppenspezifische Bedürfnisse, sind auch diese aus dem quotierten Sample gut ableitbar. Von der Menge her stellen zumindest zehn Personen je Teilsamplegruppe eine (wahrscheinlich noch) ausreichende qualitative Erhebungsbasis dar.

Zusammenfassung möglicher Vorgehensweise(n):

- Quotenauswahl [68FF.], Screening über Screening- bzw. Sondierungsfragen zur Identifizierung der Befragteilgruppen ist dabei empfehlenswert [122].
- Je nach Offenheit der Gesprächsführung Gesprächsleitfaden [113] oder qualitativer Fragebogen [115].
- Die Ergebnis-Auswertung und -Darstellung erfolgen nach qualitativen Grundsätzen [134FF. und 149FF.].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Setting, Stichprobe und Ergebnisinterpretation: Die Personen, die für die Einzelinterviews willkürlich ausgewählt wurden, sind Individuen mit wahrscheinlich divergierenden Meinungen. Bei der Ergebnisinterpretation werden diese Einzelmeinungen fälschlich generalisiert: Es erfolgt ein Schluss auf das Verhalten ALLER Studierenden: Meinen z.B. zwei Drittel der Gesprächskontakte, Sachbücher müssten unbedingt im Format A5 erhältlich sein, wird daraus abgeleitet, dass 66% aller Studierenden das Buch in A5 kaufen würden. → Die hier angewandte Methodik kann ausschließlich dazu herangezogen werden, ein für möglichst viele ideales Fachbuch (verbalisierend) zu beschreiben. In einer ganzheitlichen Sichtweise können Rückschlüsse auf die konkrete Buchgestaltung abgeleitet werden. Aus der offensichtlichen Wichtigkeit von A5 auf ein (prozentuelles) Kaufverhalten zu schließen, wäre völlig unzulässig. Mit qualitativen Settings lassen sich keinesfalls zahlenmäßige Verhältnisse in Grundgesamtheiten abbilden.

... zwischen Setting und technischer Auswertung: Trotz qualitativen Settings werden mehr oder weniger ausschließlich geschlossene Fragen gestellt, auf offene Fragestellungen wird großteils verzichtet. Bei der Auswertung sind somit fast nur prozentuelle Häufigkeiten ermittelbar. → Da Generalisierungen bei qualitativen Erhebungen mit fast immer kleinen Samples nicht erfolgen dürfen, machen Prozentauswertungen hier wenig Sinn.

9.2 | Abtestung von Kommunikationskonzepten

Eine deutsche Firma entwirft mehrere Kommunikationskonzepte für die Zielgruppe der Seniorinnen und Senioren im Alter von 60 bis 75 Jahren. Die Entwürfe sollen in Deutschland möglichst rasch bundesweit auf ihre Stimmigkeit, Akzeptanz und Kommunikationsleistung überprüft werden.

Dazu könnten in einer süddeutschen, west-, ost- und norddeutschen Hauptstadt jeweils eine, besser jeweils zwei, Gruppendiskussion(en) durchgeführt werden. Jede Gruppe setzt sich zu gleichen Teilen aus Personen der Zielgruppe zusammen, die teils IN der Stadt, teils außerhalb wohnen.

In jeder Fokusgruppe werden die Kommunikationskonzepte rund eine bis zwei Stunden lang diskutiert. Besonderes Augenmerk liegt dabei darauf, wie gut die Entwürfe aus Zielgruppensicht zur Firma passen und beurteilt werden.

Die Ergebnisse aller Diskussionen liegen bereits nach ein bis zwei Wochen in einem schriftlichen Bericht vor.

Zusammenfassung möglicher Vorgehensweise(n):

- Willkürliche Auswahl [68] aus der Testpersonen-Datenbank eines Instituts oder Interessentensuche in städtischen Fußgängerzonen der entsprechenden Regionen. Pro Gruppendiskussion werden zehn bis zwölf Personen rekrutiert, damit eine Nettoanwesenheit von zumindest acht Teilnehmenden erreicht wird.
- Diskussionsleitfaden [114].
- Die Ergebnis-Auswertung und -Darstellung erfolgen nach qualitativen Grundsätzen [134^{FF.} und 149^{FF.}].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Setting, Stichprobe und Ergebnisinterpretation: Die beiden Gruppen, die sich in Süddeutschland zur Diskussion treffen, setzen sich zufälligerweise aus Einzelpersonen mit stark divergierenden Meinungen zusammen. Die Ergebnisinterpretation generalisiert stark aus diesen Einzelmeinungen und schließt auf das Verhalten aller Seniorinnen und Senioren des Zielgebiets. Dabei wird sogar die prozentuelle Meinungsverteilung in den beiden Gruppen auf die Grundgesamtheit übertragen. → Dieses Vorgehen ist methodisch nicht korrekt: Qualitative Methodik dieser Art ist dazu geeignet, das Pro und Kontra der Kommunikationskonzepte verbalisierend zu beschreiben. Daraus können Rückschlüsse auf die konkrete Kommunikationsgestaltung abgeleitet werden. Zahlenmäßige Verhältnisse in der Grundgesamtheit lassen sich an Fokusgruppen jedoch keinesfalls festmachen.

... zwischen Leitfaden und Auswertung: Die Erhebung hat unter anderem auch das Ziel, herauszufinden, wie kaufanregend die einzelnen Kommunikationskonzepte erlebt werden. Die beiden Gruppen im Norden haben sehr intensiv diskutiert, die Diskussionsleitung konnte das Zeitbudget kaum einhalten. Der Aspekt „kaufanregend“ wurde deshalb nur mehr kurz am Rande besprochen. Im Ergebnisbericht lässt sich dieses Kriterium deshalb nur aus jenen Gruppen gut ableiten, bei denen es auch „ordentlich“ zur Sprache kommen konnte. → Wenn im Diskussionsablauf einzelner Gruppen inhaltliche Details fehlen, kann sich das verzerrend auf das Gesamtergebnisbild auswirken. Nicht vorhandene „Meinungsteile“ könnten ja völlig andere sein als die in der Summary zusammengefassten.

9.3 | Analyse der Käuferinnen- und Käufer-Struktur veräußerter Artikel

Ein Versandhändler möchte die soziodemografische Struktur und Kaufgewohnheiten seiner Kundinnen und Kunden kennenlernen. Dazu analysiert er alle Verkaufsvorgänge im Webshop im Verlauf der letzten sechs Monate.

Hier bietet sich zunächst eine Totalerhebung an. Verkaufsvorgänge können die Forschungsteilnahme nicht verweigern und damit die Ergebnisdaten nicht strukturell verzerren. Damit ist das Ergebnis völlig „repräsentativ“, weil es ja alle Fälle abdeckt. Es gibt auch keine Schwankungsbreiten zu berücksichtigen oder Signifikanztests zu berechnen: Da ALLE Verkäufe analysiert werden, müssen keine Stichprobenergebnisse auf eine dahinterstehende Grundgesamtheit projiziert werden.

Ist die Totalerhebung wegen hoher Datenmengen (z.B. hunderttausende Verkäufe) nicht durchführbar, muss eine Stichprobe Abhilfe schaffen. Als Setting erster Wahl kann zunächst eine reine Zufallsauswahl von z.B. 5.000 Verkaufsvorgängen angedacht werden. Denn obwohl bei dieser Größenordnung die Datenmenge für die Auswertung überschaubar bleibt, bewegen sich die statistischen Schwankungsbreiten noch im Rahmen.

Bei derartigen reinen Zufallsauswahlen entsteht jedoch möglicherweise ein Problem: Selten verkaufte Artikel wie z.B. (teure) Spezial-Produkte können der reinen Zufälligkeit der Auswahl zum Opfer fallen. Um dem entgegenzuwirken, müsste vor der zufälligen Ermittlung der Verkaufsdatensätze eine Vorschichtung selten und öfter gewählter Produktgruppen erfolgen. Erst aus diesen beiden Erhebungsmengen erfolgt dann die endgültige Datenauswahl für die Analyse.

Da es sich bei diesem Vorhaben um eine Analyse vorhandener Daten handelt, gibt es keinen Fragebogen. Die Auswertungsmöglichkeiten sind aber dennoch von den Messniveaus der im Webshop erfassten Merkmale abhängig. Die Erhebung kann sich hier deshalb wahrscheinlich nicht (mehr) primär nach Forschungsfragen richten: Erkenntnisinteressen müssen auch nach den Möglichkeiten ihrer Beantwortbarkeit definiert werden.

Zusammenfassung möglicher Vorgehensweise(n):

- Vollerhebung [43F.] ODER einfache zufällige Stichprobenauswahl aus der Datenbank [63F.] ODER vorgeschichtete Zufallsauswahl (kleine Untergruppen disproportional verstärkt, Faktorengewichtung auf repräsentativ [64FF.]). Die Größe der Stichprobe richtet sich nach der maximalen Höhe der Schwankungsbreiten, die noch akzeptiert werden kann [88FF.].
- Kein Fragebogen, nur Auswertungsplan; Forschungsfragen richten sich nach den Möglichkeiten der vorhandenen Datenstrukturen.
- Die Ergebnis-Auswertung und -Darstellung erfolgen nach quantitativen Grundsätzen [136FF. und 149FF.]. Wurden die Datensätze zufällig ausgewählt, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Forschungsfragen und technischer Auswertung: Bei der Definition der konkreten Erkenntnisinteressen bzw. Forschungsfragen ist unbedingt an die spätere Auswertbarkeit der Daten zu denken. Merkmale, die in der Datenstruktur des Webshops nicht vorliegen, dürfen nicht zur Beantwortung einer Forschungsfrage notwendig gemacht werden. → Liegt z.B. die „Besuchshäufigkeit“ des Webshops nur in der Form vor, an welchem Tag (Datum) der letzte Besuch erfolgte, könnte eine Forschungsfrage zur durchschnittlichen Besuchsfrequenz von Einzelpersonen nicht passgenau beantwortet werden.

9.4 | Inhaltsanalyse (zehn Jahre) über Printmedien-Berichterstattung

Eine Inhaltsanalyse vergleicht die Berichterstattung der reichweitenstärksten tagesaktuellen Printmedien – drei Boulevard-Medien und drei Qualitätszeitungen – eines Landes in den letzten zehn Jahren zu einem bestimmten Thema.

In einer zentralen Datenbank, in der ALLE Zeitungsartikel der bezeichneten Medien (auch für den Untersuchungszeitraum archiviert aufzufinden sind, lassen sich mithilfe einer Schlagwortsuche insgesamt 1.000 Artikel identifizieren, die thematische Relevanz besitzen. Davon fallen 100 Artikel auf die Boulevard-, 900 auf die Qualitätspresse.

Da jeder Artikel intensiv inhaltsanalytisch betrachtet werden soll, würde eine Totalerhebung der 1.000 Artikel den Rahmen vorhandener Personalressourcen sprengen. Durchführbar erscheint eine Analyse von höchstens 400 Artikeln: Zu ziehen ist also eine Stichprobe im Umfang von 40% der Grundgesamtheit der 1.000 identifizierten Artikel.

Dabei taucht ein Problem auf: 40% der 900 Qualitätspresse-Artikel stellen eine ausreichende Menge für eine empirische Analyse dar. 40% der insgesamt nur 100 Boulevard-Artikel wären aber nur 40 Artikel, die in die Erhebung gelangen: Das ist für statistisch abgesicherte und inhaltlich generalisierbare Aussagen deutlich zu wenig.

Die Studienverantwortlichen entscheiden sich deshalb für eine Vorschichtung, disproportionale Stichprobe und Ergebnismengewichtung: Die 100 Boulevard-Artikel werden, um statistisch höherwertige Aussagen zu erzielen, in ihrer Gesamtheit total erhoben. Aus den relevanten Artikeln der Qualitätspresse hingegen erfolgt eine Zufallsauswahl: Jeder dritte (der zuvor zufällig sortierten) Artikel (= 300 der 900 Artikel) gelangt in die Analyse. Die endgültige Analysemenge enthält damit die forschungstechnisch „durchführbaren“ 400 Artikel.

Nach Abschluss der Inhaltsanalyse können nun sinnvolle Aussagen über die Qualitäts- UND Boulevard-Artikel, jedoch nur IN ISOLIERTER BETRACHTUNG getätigt werden.

Für eine Ergebnis-GESAMTaussage ist unbedingt wieder eine Abgewichtung der Boulevardpresse und Aufgewichtung der Qualitätsmedien erforderlich: Das ursprüngliche Mengenverhältnis in der Grundgesamtheit war ja 100 : 900 Artikel. Die Artikel aus der Boulevardpresse haben also 10% der Grundgesamtheit umfasst. In der gezogenen Stichprobe ist das Verhältnis jetzt jedoch 100 : 300. Der Boulevardteil hat damit einen (deutlich zu hohen) Mengenanteil von 25%. Um die ursprüngliche Verhältnismäßigkeit wiederherzustellen, müssen die Boulevardartikel deshalb mit einem Faktor von 0,4 abgewichtet werden. Damit werden aus den 100 erhobenen Artikeln rechnerische 40, also die erforderlichen 10% der 400er-Gesamtstichprobe. Aus den 300 Artikeln der Qualitätspresse werden mit einem Gewichtungsfaktor von 1,2 rechnerische 360 Artikel. Damit umfasst die Stichprobe auch nach der Gewichtung noch 400 Artikel, jetzt aber wieder im richtigen Mengenverhältnis von 10 : 90%.

Die Forschungsfragen an die Erhebung müssen mittels Codebuch und Codierschema umgesetzt (= operationalisiert) werden.

Zusammenfassung möglicher Vorgehensweise(n):

- Vorgeschichtete Zufallsauswahl (kleine Untergruppe als Vollerhebung, Faktorengewichtung auf repräsentativ [64FF.]). Die Größe der Stichprobe richtet sich nach der maximalen Höhe der Schwankungsbreiten, die noch akzeptiert werden kann [88FF.].

- Codebuch und Codierschema zur Datenerfassung [25FF.]. Enthält das Codierschema qualitative Elemente, ist eine (induktive oder deduktive) Kategorienbildung anzudenken [134FF.].
- Die Ergebnis-Auswertung und -Darstellung erfolgen vorwiegend nach quantitativen Grundsätzen [136FF. und 149FF.]. Da die Daten der Qualitätspresse auf einer Zufallsstichprobe beruhen, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.]. Die Boulevarddaten wurden total erhoben, ihre Ergebnisse besitzen somit KEINE statistische Unschärfe. Alle qualitativen Elemente des Codierschemas sollten mit Zitation erläuternder Textbeispiele ergänzt werden [150].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Forschungsfragen, Codierschema und Auswertung: Bei der Definition der konkreten Erkenntnisinteressen (Forschungsfragen) muss unbedingt an deren Zusammenspiel mit dem Codierschema gedacht werden. So können etwa Merkmale, die im Codebuch nur nominal definiert werden, später keiner metrischen Auswertung zugeführt werden. → Beispielsweise soll die Intensität der Verwendung einzelner Begriffe gemessen werden. Codiert wird aber nur mit „ja | nein“, ob die einzelnen Begriffe überhaupt vorkommen. Damit wird eine Aussage über Intensität unmöglich. Dazu wäre eine mengenmäßige Abbildung der Häufigkeiten (in Form einer Skalierung, konkreten Zählung o.Ä.) notwendig gewesen.

... zwischen Grundgesamtheit, Stichprobe und Ergebnissen: Bei der Auswahl der Artikel werden der Einfachheit halber 100 Artikel der Boulevard- und 100 Artikel der Qualitätspresse analysiert und gemeinsam ausgewertet. Das Ergebnis wird generalisiert. → Die dabei entstehenden Aussagen sind keinesfalls repräsentativ. In der Grundgesamtheit ist das Verhältnis zwischen Boulevard- und Qualitätsmedien 1 : 9. In der hier angewendeten Stichprobe hingegen 1 : 1.

9.5 | Inhaltsanalyse (sechs Monate) der Chronik-Artikel einer Tageszeitung

Alle im Verlauf der letzten sechs Monate publizierten Artikel im Chronikteil der (fiktiven) Zeitung „Die Tageszeitung“ werden einer Inhaltsanalyse unterzogen.

Dazu scheint zunächst eine Vollerhebung am besten geeignet zu sein. Zeitungsartikel verweigern keine Forschungsteilnahme, die Ergebnisdaten können somit nicht strukturell verzerrt sein. Damit ist das Ergebnis völlig repräsentativ, weil es ja alle Fälle abdeckt. Es gibt auch keine Schwankungsbreiten zu berücksichtigen oder Signifikanztests zu berechnen: Da ALLE Artikel analysiert werden, müssen keine Projektionen von Stichprobenergebnissen auf dahinterstehende Grundgesamtheiten erfolgen.

Ist die Anzahl der Artikel zu groß oder eine Vollerhebung zu aufwendig, bietet sich bei diesem Forschungsvorhaben ein reines Zufallssample an. Die Grundgesamtheit stellen die an allen Tagen der sechs Beobachtungsmomente im Chronikteil der „Tageszeitung“ publizierten Artikel dar. Zunächst erfolgt eine Auflistung aller Erscheinungstage der „Die Tageszeitung“. Daraus werden per Zufall Tage ausgewählt. Von diesen Tagen wird – wieder zufällig – jeweils einer der am jeweiligen Tag publizierten Chronik-Artikel in die Analyse

miteinbezogen. Das könnte z.B. alternierend der erste und letzte Artikel aus dem Chronikteil des jeweiligen Tages sein. Oder ALLE Chronik-Artikel der zufällig ausgewählten Tage gelangen in die Analyse.

Sind unter der Woche und am Wochenende unterschiedliche Artikelmengen zu erwarten, ist eventuell eine zusätzliche Verfeinerung der Stichprobe in Betracht zu ziehen: Die Grundgesamtheit könnte in diesem Fall noch nach Tagestypen vorgeschichtet werden. Die Anteile der ausgewählten Tage von z.B. Mo bis Fr (unter der Woche) sowie Sa und So (Wochenendausgaben) an der Gesamtstichprobe würden dann dem Verhältnis dieser Tage in der Grundgesamtheit entsprechen.

Die Forschungsfragen an die Erhebung müssen mittels Codebuch exakt definiert (= operationalisiert) und in einem Codierschema für die Datenerfassung umgesetzt werden.

Zusammenfassung möglicher Vorgehensweise(n):

- Vollerhebung [43F.] ODER einfache zufällige Stichprobenauswahl [63F.] ODER vorgeschichtete Zufallsauswahl (nach Tagestypen) [64F.]. Die Größe der Stichprobe richtet sich nach der maximalen Höhe der Schwankungsbreiten, die noch akzeptiert werden kann [88FF.].
- Codebuch und Codierschema zur Datenerfassung [25FF.]. Enthält das Codierschema qualitative Elemente, ist eine (induktive oder deduktive) Kategorienbildung anzudenken [134FF.].
- Die Ergebnis-Auswertung und -Darstellung erfolgen – den Forschungsfragen entsprechend – nach qualitativen [134FF.] oder quantitativen Grundsätzen [136FF. und 149FF.]. Beruhen die Daten auf einer Zufallsstichprobe, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.]. Eventuell vorhandene qualitative Elemente des Codierschemas sollten mit Zitation erläutern-der Textbeispiele ergänzt werden [150].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Forschungsfragen, Codierschema und Auswertung: Die konkreten Forschungsfragen müssen unbedingt auch im Codierschema mit den ihnen entsprechenden Messniveaus „abgebildet“ werden. → Ist z.B. eine Ergebnisdarstellung nach Artikellänge Teil der Erkenntnisinteressen, muss dementsprechend eine diesbezügliche Variable mitaufgenommen werden.

9.6 | Inhaltsanalyse der gesamten Mitarbeiterkommunikation

Ein Unternehmen überprüft die gesamte Mitarbeiterkommunikation (Rund-E-Mails, Briefe, Mitarbeiterzeitungen, Intranet usw.) der letzten zwölf Monate inhaltsanalytisch.

Auf den ersten Blick bietet sich hier wieder eine Vollerhebung an. Auf Kommunikationsinhalte kann wohl in der Regel zumindest ein Jahr lang zugegriffen werden. Kommunizierte Beiträge können keine Erhebung verweigern, die Analyse kann also keine strukturell verzerrten Ergebnisse liefern. Werden alle Fälle analysiert, entfallen als weiterer Vorteil auch jegliche Stichprobenfehler.

Ist eine Totalerhebung zu aufwendig, könnten aus den letzten 52 Kalenderwochen z.B. zehn zufällig ausgewählt werden. Die Gesamtkommunikation dieser zehn Wochen gelangt in die Analyse. Das reduziert die Menge der Totalerhebung auf ein Fünftel.

Sind bei den Erkenntnisinteressen der Erhebung Saisonalitäten relevant, wäre eine diesbezügliche Vorschichtung sinnvoll: Angenommen, im Unternehmen fallen z.B. zwei Drittel des Jahres auf Wochen mit relativ konstanten Mitarbeiterzahlen. Im restlichen Drittel herrscht starke Personalfuktuation vor. In Zeiten umfangreicher Personalveränderungen findet erfahrungsgemäß mehr interne Kommunikation statt. Die ausgewählten Analysewochen verteilen sich deshalb ebenfalls in diesem Verhältnis auf die beiden Zeittypen.

Eine in diesem Beispiel zusätzlich sehr wichtige Festlegung betrifft die Codiereinheiten: E-Mails, Briefe, Zeitungen und Intranet stellen je „Ausgabe“ von der Themenmenge her völlig unterschiedliche Dimensionen dar. Überwiegen bei manchen Kanälen pro Kommunikationsfall einzelne Themen, handelt es sich an anderer Stelle um eine Vielfalt kommunizierter Inhalte. Eine Möglichkeit, diesem Umstand Rechnung zu tragen, besteht darin, die (Auswahl- und) Codiereinheiten als „thematische Einheiten“ zu definieren.

Die Forschungsfragen an die Erhebung müssen mittels Codebuch und Codierschema umgesetzt (= operationalisiert) werden.

Zusammenfassung möglicher Vorgehensweise(n):

- Bevorzugt Vollerhebung [43F.] aus allen oder den Inhalten zufällig ausgewählter Kalenderwochen [63F.]. Eventuelle Vorschichtung der Kalenderwochen nach Saisonalitäten [64F.]. Die Größe der Stichprobe richtet sich nach der maximalen Höhe der Schwankungsbreiten, die noch akzeptiert werden kann [88FF.].
- Codebuch und Codierschema zur Datenerfassung [25FF.]. Enthält das Codierschema qualitative Elemente, ist eine (induktive oder deduktive) Kategorienbildung anzudenken [134FF.].
- Die Ergebnis-Auswertung und -Darstellung erfolgen – den Forschungsfragen entsprechend – nach qualitativen [134FF.] oder quantitativen Grundsätzen [136FF. und 149FF.]. Beruhen die Daten auf einer Zufallsstichprobe, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.]. Eventuell vorhandene qualitative Elemente des Codierschemas sollten mit Zitation erläuternder Textbeispiele ergänzt werden [150].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Forschungsfragen, Codierschema und Auswertung: Bei der Definition der konkreten Erkenntnisinteressen (Forschungsfragen) muss unbedingt an deren Zusammenspiel mit dem Codierschema gedacht werden. So können etwa Merkmale, die im Codebuch nur mit einzelnen Ausprägungen definiert werden, später keiner Auswertung nach mehrfachem Vorkommen (Mehrfachauswertung) zugeführt werden. → Beispielsweise soll analysiert werden, ob in denselben thematischen Einheiten auch mehrere Führungskräfte (und wenn, welcher Ebene) zu Wort kommen. Erfasst und codiert wird dazu aber nur EINE Variable FÜHRUNGSKRAFT. Somit ist bei Vorkommen MEHRERER Personen weder eine Aussage über die Anzahl noch über deren Position im Unternehmen möglich. Dazu wären mehrere Variablen notwendig gewesen.

... **zwischen Grundgesamtheit, Stichprobe und Ergebnissen:** In die Auswahl gelangen jeweils 20 thematische Einheiten aus E-Mails, Briefen, Mitarbeiterzeitungen und Intranet-Beiträgen. Das Ergebnis wird generalisiert. → Die dabei entstehenden Aussagen sind wahrscheinlich nicht repräsentativ. In der Grundgesamtheit ist das Verhältnis der Anzahl thematischer Einheiten zwischen den Kommunikationskanälen vermutlich nicht ausgeglichen. Für ein repräsentatives Ergebnisbild ist es erforderlich, die Auswahl der Analyseelemente an ihrer mengenmäßigen Verteilung in der Grundgesamtheit zu orientieren.

9.7 | Mystery-Shopping eines Lebensmitteldiskonters

Ein deutscher Lebensmitteldiskonter plant in seinen Filialen ein Mystery-Shopping.

Das wohl einfachste Setting wäre hier eine Vollerhebung aller Filialen.

Ist eine Vollerhebung aller Filialen nicht durchführbar, bietet sich eine Zufallsstichprobe aus dem gesamten Filialnetz an. Um ein repräsentatives Bild zu erhalten, müssen Strukturmerkmale definiert werden, deren Verteilung in der Stichprobe auf Übereinstimmung mit der Grundgesamtheit überprüft werden können. Das können bei Mystery-Tests Filialstandort (Bundesland) oder Filialgröße (Mitarbeitende, Umsatz oder Verkaufsvorgänge) sein. Die durchzuführenden Mystery-Tests müssen derart auf das Filialnetz aufgesplittet werden, dass sie dessen Verteilung innerhalb der definierten Strukturmerkmale widerspiegeln. Vereinfacht gesagt müssen große Filialen öfter getestet werden als kleine. Um Repräsentativität sicherzustellen sollte die Zufallsstichprobe anhand der Strukturmerkmale vorgeschichtet werden.

Eine Quotenstichprobe der Filialen (z.B. 100 Tests in Bayern, 100 in Hessen usw.) wäre hier wohl möglich, stellt aber wegen der zu befürchtenden subjektiven Auswahl der Testorte durch die Testenden eine weniger empfehlenswerte Variante dar. Wenn die Entscheidung trotzdem auf eine Quotenstichprobe fällt, sollten die Testpersonen nur wenig Entscheidungsspielraum bei der Filialauswahl zugestanden bekommen. Jedenfalls müssen auch die Quoten unbedingt die definierten Strukturmerkmale widerspiegeln.

Unabhängig von der Art der Stichprobe ist auf eine möglichst breite Verteilung der Tests über Tage, Wochentage und Tageszeiten zu achten. Als Tagestypen und Uhrzeiten ließe sich definieren: Montag, Dienstag bis Donnerstag, Freitag, Samstag bzw. 7:00 bis 9:00, 9:01 bis 11:00, 11:01 bis 14:00, 14:01 bis 16:00, 16:01 bis 19:00 Uhr. Eine Quotierung der Filialbesuche nach diesen Tagen und Zeiten wäre eine praktikable Lösung. Im Idealfall erstreckt sich die gesamte Erhebung über mehrere Wochen.

Die Forschungsfragen an die Erhebung müssen mittels Mystery-Protokollbogen objektiv umgesetzt (= operationalisiert) werden. Mystery-Tests stellen keine Kundenbefragungen dar, sondern OBJEKTIVE Bestandsaufnahmen. Die Testpersonen dürfen deshalb nur objektive Gegebenheiten erfassen, ihre eigene Meinung ist wenn, dann höchstens ergänzend, am Rande gefragt.

Zusammenfassung möglicher Vorgehensweise(n):

- Vollerhebung [43f.] ODER einfache zufällige Stichprobenauswahl aus dem Filialnetz [63f.] ODER vorgeschichtete Zufallsauswahl (nach zu definierenden Strukturmerkma-

len, kleine Untergruppen disproportional verstärkt, Faktorengewichtung auf repräsentativ [64FF.]). Die Größe der Stichprobe (= Mystery-Tests) richtet sich nach der maximalen Höhe der Schwankungsbreiten, die noch akzeptiert werden kann [88FF.].

- Mystery-Protokollbogen [32F.] zur Datenerfassung.
- Die Ergebnis-Auswertung und -Darstellung erfolgen hier nach quantitativen Grundsätzen [136FF. und 149FF.]. Wurden die Datensätze zufällig ausgewählt, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.].

Mögliche Unstimmigkeiten könnten auftreten ...

... **zwischen Grundgesamtheit, Stichprobe und Ergebnissen:** Für die Stichprobe werden nur große Filialen ausgewählt. Dort ist das Kundenaufkommen am größten und eher mit Defiziten zu rechnen – so die Annahme der Forschungsleitung. Außerdem wird den Testpersonen freigestellt, wann sie den Test machen: Einkäufe finden ja auch zu x-beliebigen Zeiten statt. → Die Beschränkung auf große Filialen wegen ausschließlich dort zu erwartender Probleme stellt ein Präjudiz dar. Vielleicht stellen sich gerade kleine Filialen problembehafteter dar. Wird darüber hinaus die Zeitlage der Besuche freigestellt, fehlt jegliche Kontrolle über die Verteilung der Tests. Etwa werden nebenberuflich tätige Testpersonen ihre Tests bevorzugt nach der Arbeit, am späteren Nachmittag durchführen. Vielleicht treten die größten Probleme aber am Vormittag auf ...? Ohne repräsentative Aufteilung der Mystery-Checks auf Filialen und Zeitlagen ist deshalb mit einem (deutlich) verzerrten Ergebnisbild zu rechnen.

... **zwischen Erkenntnisinteressen und Erhebungsprotokoll:** Im Erhebungsprotokoll wird gefragt: „Wie freundlich war der/die Mitarbeiter/in, den/die Sie am Regal nach der Schuhcreme gefragt haben? Urteilen Sie bitte mit Schulnoten von 1 bis 6“. → Mystery-Tests müssen streng objektiviert ablaufen. Freundlichkeit ist für jeden Menschen (und damit auch die Testenden) anders definiert. Die Ergebnisse einer Bewertung dürfen aber nicht von der Testperson abhängen. Es müssen Indikatoren gefunden werden, die „Freundlichkeit“ kennzeichnen. Ein derartig objektiver Gradmesser könnte im Beispiel sein: „Wie hat der/die Mitarbeiter/in auf Ihre Frage nach der Schuhcreme reagiert? >>> hat mich auf die Abteilung verwiesen | hat mir den Platz mit ‚dort drüben‘ gezeigt | ist mit mir zum entsprechenden Regal gegangen | hat anders reagiert, und zwar ...“.

9.8 | Kundenzufriedenheitsanalyse eines Einkaufszentrums

Eine Kundenzufriedenheitsanalyse fokussiert auf weibliche und männliche Privatkundinnen und -kunden eines Einkaufszentrums in einer mittleren Stadt.

In einem Einkaufszentrum muss sich niemand beim Einkauf namentlich registrieren. Es gibt hier also keine zugängliche, bekannte Grundgesamtheit. Auch in einer repräsentativen Bevölkerungstichprobe werden kaum genügend Kundinnen und Kunden dieses Einkaufszentrums zu finden sein (= kleine Inzidenz). Außerdem ist das Einzugsgebiet nicht bekannt, also von wie weit die Kundschaft zufährt. Damit fehlt die Information, wie breit eine Bevölkerungstichprobe geografisch auszudehnen wäre, um ein repräsentatives Kundenbild zu bekommen.

In einem derartigen Fall besteht die einzige Möglichkeit wohl darin, auf eine (willkürliche) Auswahl direkt vor Ort zurückzugreifen. Die Auswahl der Befragten sollte auf jeden Fall zeitlich sehr breit gestreut (nach Uhrzeit und Wochentagen) und über einen längeren Zeitraum erfolgen. Damit ist die Wahrscheinlichkeit größer, die unterschiedlichen Kundengruppen in ihrer Gesamtheit möglichst gut abzudecken. Auch auf das Abbilden (oder Vermeiden – je nach Erkenntnisinteresse) spezieller saisonaler Unregelmäßigkeiten in der Kundenstruktur sollte geachtet werden. Zufälligkeit bei der Stichprobenziehung lässt sich durch Ansprechen jeder fünften, zehnten, 20. (o.ä.) Person erreichen, die einen bestimmten Markierungspunkt überschreitet.

Soll sichergestellt sein, dass das Meinungsbild wichtiger Teilsegmente nicht fehlt, kann hier (zusätzlich) auch ein Quotensample zum Einsatz kommen: Das würde z.B. eine Mengenfestlegung nach Geschlecht, Altersgruppe und Einzugsgebiet treffen. Die Teilgruppengrößen orientieren sich an eventuell vorhandenem Wissen über deren Verteilung in der Grundgesamtheit oder werden alternativ frei festgelegt.

Vollständige Repräsentativität wird hier unerreichbar sein: Einerseits können unbekannte Effekte die Käuferstrukturen verändern (z.B. Verkaufsaktionen einzelner Firmen). Andererseits ist mangelnde Information über die strukturelle Zusammensetzung der Grundgesamtheit anzunehmen: Damit ist nicht beurteilbar, ob die Stichprobenstruktur jener der Grundgesamtheit entspricht. Die angeführten Maßnahmen erlauben es aber, sich dem Idealfall der Repräsentativität zumindest weitestmöglich annähern zu können.

Bei Erhebungen dieser Art kommen in der Regel quantitative Fragebögen zum Einsatz.

Zusammenfassung möglicher Vorgehensweise(n):

- Willkürliche Auswahl [68] oder Quotenstichprobe [68FF.] nach zu definierenden Strukturmerkmalen, die erhebungsrelevant sind. Zur besseren Objektivierung der Auswahl zufällige Streuung der Erhebungstage, Tageszeiten und Zielpersonen [73FF.]. Repräsentativität [55FF.] kann hier kaum gemessen bzw. gewährleistet werden.
- Wenn die Auswahl keine Zufallskomponenten enthält, ist die Berechnung von Schwankungsbreiten [76FF.] eigentlich unzulässig (erfolgt aber dennoch oft als „Richtwert“ der Ergebnisgüte).
- Meist quantitativer Fragebogen [Beispiel 101F., im Detail 116FF.], der vereinzelt auch qualitative Elemente [123FF.] beinhalten kann.
- Die Ergebnis-Auswertung und -Darstellung erfolgen hier wohl überwiegend nach quantitativen Grundsätzen [136FF. und 149FF.] mit eventuell eingestreuten qualitativen Elementen [134FF.]. Wurden die Datensätze zufällig ausgewählt, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Stichprobe und Ergebnisinterpretation: Bei der Befragung im Einkaufszentrum wählen die Erhebungspersonen an zwei aufeinanderfolgenden Samstagen ab 10 Uhr vormittags bei der zentralen Eingangstür willkürlich 100 Männer und 100 Frauen aus: Sie laden jede ihnen „ansprechbar“ erscheinende Person zur Befragung ein, bis sie jeweils 100 Interviews gesammelt haben. → Damit nähert sich das Sample keinesfalls auch nur

irgendwie der Repräsentativität an: Zumindest fehlt eine Streuung der Befragungen nach Wochentagen, Tageszeiten oder Einzugsbereich der Kundinnen und Kunden. Von der Momentaufnahme zweier strukturell ähnlicher Vormittage auf eine Grundgesamtheit zu schließen ist methodisch unzulässig. Zusätzlich verzerrt auch noch der Sympathiefaktor zwischen Befragten und Befragenden, ohne den ein Interview gar nicht zustandekommt, das Ergebnis.

9.9 | Mitarbeiterbefragung einer Möbelkette

Eine deutsche Möbelkette plant, ihre Mitarbeiterinnen und Mitarbeiter zu befragen.

Dazu werden als Grundgesamtheit definiert: Alle Personen, die in einer der Filialen oder der Verwaltung seit mindestens einem halben Jahr im Ausmaß von mindestens 20 Wochenstunden fix angestellt sind und keiner Führungsebene (auch nicht Teamleitung) angehören.

Hier kommt – bei halbwegs überschaubarer Anzahl Mitarbeitender – durchaus eine Vollerhebung in Frage. Beschäftigte sind üblicherweise in Unternehmensdatenbanken erfasst, also direkt kontaktierbar. Im Falle einer Gesamterhebung ist aber unbedingt auf Folgendes zu achten: Da von Teilnahmeverweigerungen auszugehen ist, muss der Rücklauf der Interviews (noch) die Grundgesamtheit repräsentieren. Um das beurteilen zu können, müssen (sinnvolle) Strukturmerkmale definiert und festgelegt werden, deren Verteilung in der gesamten Belegschaft bekannt ist. Nur dann kann die Struktur jener, die sich an der Erhebung beteiligt haben, auf ihre Übereinstimmung mit der Grundgesamtheit (= Repräsentativität) überprüft werden. Das könnten im gegenständlichen Fall Filialstandort (Bundesland), Filialgröße (Anzahl der Mitarbeitenden, Umsatz oder Verkaufsvorgänge), Beschäftigtenart (Arbeitende oder Angestellte) oder auch Abteilung (Filiale oder zentrale Verwaltung, und dort Vertrieb, Einkauf, Marketing und Werbung usw.) sein.

Nach derartigen Kriterien wäre auch eine – hier ebenfalls anwendbare – Zufallsstichprobe vorschichtbar. Für jede Teilgruppe dieser Vorschichtung wird zunächst ihre anteilmäßig erforderliche Größe ermittelt. Erst dann erfolgt – wieder je Teilgruppe – eine Zufallsauswahl aus den Mitarbeiterdaten. Zu kleine Gruppen werden aufgestockt, also mengenmäßig vergrößert (und bei der Gesamtergebnisermittlung wieder abgewichtet).

Derartige Zufallsstichproben liefern wahrscheinlich bei einer Vielzahl von Mitarbeiterbefragungen repräsentativere Ergebnisse als eine Vollerhebung. Durch die Vorschichtung wird es möglich, ein besseres Abbild der Grundgesamtheit zu zeichnen als bei einer Totalerhebung: Dort kann der oft unterschiedlich starke Interview-Rücklauf aus den verschiedenen Firmenbereichen das Ergebnis rasch verzerren (z.B. befragungsaффineres Büropersonal versus ablehnenderes Verkaufspersonal).

Bei Mitarbeiterbefragungen kommen in der Regel quantitative Fragebögen zum Einsatz. Methodisch ist es in den meisten Unternehmen möglich, mittels Online-Befragung vorzugehen. Online-Befragungen bieten den großen Vorteil, durch den Versand von Reminder-E-Mails den Rücklauf zu erhöhen. Ob die versendeten Links personalisiert sind oder nicht, ist situationsspezifisch zu entscheiden. Bei den wohl meisten Online-Erhebungen im Mitarbeiterbereich – ob personalisiert oder nicht – treten immer wieder zwei Problembereiche auf: 1. treten bei der Belegschaft Bedenken auf, dass die Erhebung nicht anonym ist;

2. sind Arbeitsbereiche, die über keine Büros mit EDV-Ausstattung verfügen, schwieriger zu erreichen als andere. Einige Mitarbeiterbefragungen gehen deshalb (auch) den – deutlich umständlicheren – Weg über eine schriftliche Befragung. Diese wird an die Wohnadressen der Mitarbeiterinnen und Mitarbeiter verschickt. Ein frankiertes Rücksendekuvert liegt bei. Das ist einer Abgabe des Fragebogens in der Firma (z.B. Einwurf in Boxen) vorzuziehen, weil dadurch der erlebte Anonymitätsgrad verbessert wird. Nachteil der schriftlichen Methode ist der höhere Zeit-, Kosten- und Organisationsaufwand sowie der zu erwartende geringe(re) Rücklauf.

Eine weitere alternative Vorgehensweise könnte den Kanal der Fragebeantwortung überhaupt freistellen: Die Mitarbeitenden erhalten einen Fragebogen mit Rücksendekuvert zugeschickt. Am Fragebogen ist auch der Link auf ein Befragungsformular aufgedruckt, das am Handy, Privat- oder Firmen-PC ebenso für die Beantwortung nutzbar ist. Ist am Papierfragebogen zusätzlich ein persönlicher Code angeführt, kann über dessen Eingabe im Online-Formular die Beantwortung sogar personalisiert erfolgen. Das wiederum erweckt aber den Anschein mangelnder Anonymität – auch wenn jeglicher Personenbezug der Daten ohnehin laut Datenschutzgesetzen ausgeschlossen werden muss.

Wird die Befragung nicht als Vollerhebung, sondern in Form einer repräsentativen Stichprobe aufgesetzt, kann als weiterer effizienter Befragungskanal das (Mobil-)Telefon ins Auge gefasst werden. Hier wird wahrscheinlich Handy bzw. Privatanschluss die bessere Alternative zum Firmenfestnetz sein.

Wie auch immer eine Mitarbeiterbefragung gestaltet ist: In jedem Fall bietet begleitende Kommunikation dabei eine wichtige Unterstützung. Diese kann vor allem auch den Aspekt der Anonymität stark in den Vordergrund rücken – und das erhöht die Ausfüllbereitschaft.

Zusammenfassung möglicher Vorgehensweise(n):

- Vollerhebung [43F.] ODER einfache zufällige Stichprobenauswahl [63F.] ODER geschichtete Zufallsauswahl (nach zu definierenden Strukturmerkmalen, kleine Untergruppen disproportional verstärkt, Faktorengewichtung auf repräsentativ [64FF.]). Die Größe der Stichprobe richtet sich nach der maximalen Höhe der Schwankungsbreiten, die noch akzeptiert werden kann [88FF.]. Bei Vollerhebungen sollte besonderes Augenmerk auf die Repräsentativität [55FF.] des Datenrücklaufs gelegt werden. Repräsentativität ist bei Online-Erhebungen nur mittels personalisierter Links überprüfbar [59FF.].
- Meist quantitativer Fragebogen [Beispiel 101F., im Detail 116FF.], der vereinzelt auch qualitative Elemente [123FF.] beinhalten kann.
- Die Ergebnis-Auswertung und -Darstellung erfolgen hier überwiegend nach quantitativen Grundsätzen [136FF. und 149FF.] mit eventuell eingestreuten qualitativen Elementen [134FF.]. Wurden die Datensätze zufällig ausgewählt, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.].

Mögliche Unstimmigkeiten könnten auftreten ...

... **zwischen Stichprobe und Ergebnisinterpretation:** Die meisten Mitarbeiterbefragungen richten sich an alle Mitarbeitenden. Die Erhebung ist abgeschlossen, von den 1.000 Mit-

arbeitenden haben 510 geantwortet, der Rücklauf liegt also bei 51%. Stolz wird ein Ergebnisbericht erstellt, das Ergebnis kommuniziert und in die Unternehmensprozesse eingesteuert. → Bei der Ergebnisermittlung erfolgt die Auswertung des Rücklaufs oft unreflektiert. Was zählt, ist die Menge an Antworten, deren Struktur wird kaum oder gar nicht kontrolliert. Das kann einen entscheidenden Fehler nach sich ziehen: Wenn sich die Struktur jener, die an der Erhebung teilgenommen haben, stark von der Grundgesamtheit unterscheidet, fehlt die Repräsentativität. Von 51% Teilnehmenden darf dann nicht ohne Weiteres auf die gesamte Belegschaft geschlossen werden. Sind die Daten spezieller Mitarbeitergruppen unterrepräsentiert, die Antworten anderer Abteilungen überrepräsentiert, führt das ziemlich sicher zu Ergebnisverzerrung(en). Hier wäre eine Ergebnisgewichtung angebracht.

... **zwischen Stichprobe und Ergebnisinterpretation:** Repräsentativitätsprobleme können auch auftreten, wenn der Datenrücklauf mit den falschen Kriterien strukturell überprüft wird. → Hinsichtlich Bundesland, Filialgröße und Abteilung passen die gewonnenen Antworten zu den Verhältnissen der Grundgesamtheit. Damit gilt Repräsentativität als bewiesen. Es könnte nun z.B. aber sein, dass die Beschäftigtenart (arbeitend oder angestellt) starken Einfluss auf die Arbeitszufriedenheit hat. Bei der Rücklaufkontrolle der Daten bleibt gerade DIESES Merkmal unbeachtet. Die Daten enthalten anteilig zu viele Meinungen von Angestellten – und DAS führt nun DOCH zu Ergebnisverzerrungen ...

9.10 | Potenzialanalyse eines Fertighausproduzenten

Ein Fertighausproduzent erhebt seine Potenziale in Österreich.

Als Hauptzielgruppe (= Grundgesamtheit) für ein neues Haus können alle jene vermutet werden, die kurz vor oder nach der Schwelle zur Familiengründung stehen oder ein gewisses Maß an Sesshaftigkeit in ihr Leben bringen wollen. Diese Menschen werden altersmäßig wohl größtenteils zwischen 20 und 45 Jahren liegen.

Als einfache Lösung ließe sich hier eine Erhebung im passenden altersmäßigen Ausschnitt einer Standard-Bevölkerungsstichprobe andenken. Z.B. werden im Rahmen einer Mehrthemenumfrage eines Instituts alle 20- bis 45-Jährigen befragt, ob sie in nächster Zeit an ein neues Haus denken. Der daraus resultierende Prozentsatz wird auf die „dahinterstehende“ absolute Anzahl der 20- bis 45-jährigen Grundgesamtheit projiziert.

Mit den (reduzierten) methodischen Möglichkeiten einer (studentischen) Einzelperson(en)gruppe) könnte hier auch eine Befragung via Social Media (im Schneeballsystem, mit darauffolgender Quotenkontrolle) angedacht werden. Um hohe Diversität der Stichprobe sicherzustellen, werden aus einer Liste von 100 interessen-, alters- und regionalspezifisch für einen Themenbezug in Frage kommenden, aber dennoch möglichst unterschiedlichen Facebook-Interessengruppen 20 zufällig ausgewählt. Die Ergebnisse aus einer derartigen Erhebung sind dann zwar nicht „beweisbar repräsentativ“, können aber zumindest als Indikatoren für ein näherungsweise themenspezifisches Meinungsbild (einer facebookaffinen Grundgesamtheit) herangezogen werden.

Absolutmengen von Menschen verschiedener Altersgruppen in der Bevölkerung finden sich in amtlichen Statistiken: Wenn 20- bis 45-Jährige in Österreich rund drei Mio. Men-

schen ausmachen und die Umfrage z.B. ein Interesse von 10% dieser Altersgruppe an Ferrih Häusern ermittelt, bedeutet das ein tatsächliches Potenzial von etwa 300.000 Personen.

Potenzialerhebungen funktionieren grundsätzlich immer nach diesem Prinzip: Die Grundgesamtheit muss größenmäßig bezifferbar sein. Liegt darüber (noch) keine verwendbare Statistik vor, ist vorab Basisrecherche notwendig. Die Ergebnisse der Erhebung, das „Potenzial“, wird anteilmäßig auf die Menge der Grundgesamtheit übertragen. Manche Erhebungen nennen das „Projektion“. Die Ermittlung von Medienreichweiten oder Zuschauerquoten von Fernsehsendungen erfolgt z.B. auf diese Art und Weise. Eines setzen alle derartigen Projektionen voraus: Sie machen nur dann Sinn, wenn die absolute Größe der Grundgesamtheit bekannt ist. Ansonsten ist nicht beurteilbar, ob z.B. 2% ein großes oder kleines Potenzial darstellen.

Bei Erhebungen dieser Art kommen quantitative Fragebögen zum Einsatz.

Zusammenfassung möglicher Vorgehensweise(n):

- Zufallsauswahl [63FF.] oder Quotenstichprobe [68FF.] nach strukturellen Merkmalen der Grundgesamtheit (verbreitet sind hier Geschlecht, Alter, Wohnort, Bildung). Die Grundgesamtheit MUSS für dieses Vorhaben größenmäßig bezifferbar sein.
- Wenn die Auswahl keine Zufallskomponenten enthält, ist die Berechnung von Schwankungsbreiten [76FF.] eigentlich unzulässig (erfolgt aber dennoch oft als „Richtwert“ der Ergebnisgüte).
- Quantitativer Fragebogen [Beispiel 101F., im Detail 116FF.], der vereinzelt auch qualitative Elemente [123FF.] beinhalten kann.
- Die Ergebnis-Auswertung und -Darstellung erfolgen hier überwiegend nach quantitativen Grundsätzen [136FF. und 149FF.] mit eventuell eingestreuten qualitativen Elementen [134FF.]. Wurden die Datensätze zufällig ausgewählt, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76FF.].

Mögliche Unstimmigkeiten könnten auftreten ...

... **zwischen Fragebogen und Erkenntnisinteresse**: Mit der Erhebung soll auch das Potenzial unter den 20- bis 30-jährigen Personen, die in Salzburg wohnen, ermittelt werden. Dazu werden die Bevölkerungsdaten der Statistik Austria abgefragt. Diese sind je Bundesland und Altersgruppe im Netz abrufbar. Im Fragebogen werden folgende Fragen gestellt: „Wo liegt der Mittelpunkt Ihrer Lebensinteressen?“ >>> „in Wien, Niederösterreich, Burgenland“ | „Oberösterreich, Salzburg“ | „Tirol, Vorarlberg“ | „Steiermark, Kärnten“ sowie „Wie alt sind Sie? Welcher der folgenden Altersgruppen ordnen Sie sich zu?“ >>> „15 bis 20“ | „21 bis 35“ | „36 bis 50 Jahre“ | „51 und älter“. → Fragebogen und Erkenntnisinteresse sind nicht deckungsgleich. Die Kategorien der Merkmale Region und Alter lassen keine Aussagen über 20- bis 30-jährige Menschen in Salzburg zu: Salzburg wurde gemeinsam mit Oberösterreich abgefragt, das Alter in der Gruppierung 21 bis 35 Jahre. Somit kann das Zielsegment aus den Daten nicht herausgelöst werden.

9.11 | Expertenbefragung kleinerer und mittlerer Unternehmen

Eine Expertenbefragung möchte Einblick in den Stellenwert des Qualitätsmanagements in kleineren und mittleren Unternehmen (KMU's) einer bestimmten Branche geben.

Für die methodische Umsetzung dieses Erkenntnisinteresses spielt die Größe der Grundgesamtheit eine entscheidende Rolle.

Ist die Branche sehr überschaubar, existieren vielleicht nur ein, zwei Handvoll kleinerer und mittlerer Unternehmen. Hier stellen wohl Leitfaden-Gespräche bzw. qualitative Interviews mittels Fragebogen die Methode erster Wahl dar. Nach telefonischem oder per E-Mail durchgeführtem Aviso finden an vereinbarten Orten Einzelinterviews mit Expertinnen und Experten der jeweiligen Firmen statt. Aus methodischer Sicht stellt dieses Vorgehen eine Vollerhebung dar. Erhebungen in derart geringem Umfang können natürlich kaum die Forderung nach Anonymität erfüllen. Große Bedeutung kommt hier der Einhaltung geltender Datenschutzbestimmungen zu. Bei der Auswertung ist deshalb die Anonymisierung von „Wer hat welche Meinung vertreten“ zu garantieren. Ein Weg dazu kann darin bestehen, zu Beginn des Berichts die Expertinnen und Experten – nach deren Zustimmung – zwar namentlich anzuführen, sie im Bericht selbst jedoch nur als „Person A“, „Person B“ usw. (ohne Verbindung zum Namen) zu zitieren.

Ist die Branche etwas größer, kann eine qualitative Methodik eventuell (noch) beibehalten werden. Die Kontaktpersonen lassen sich dann vielleicht aber bereits über eine (nach Firmengröße und/oder Region vorgeschichtete) Zufallsstichprobe ermitteln. Auch quotierte Kontaktvorgaben sind hier möglich und sinnvoll.

Sind in der Grundgesamtheit VIELE Betriebe gelistet, ist eine vorgeschichtete Zufallsstichprobe – z.B. über Region und Firmengröße – wahrscheinlich eine gute Wahl. Bei nun größerer Fallzahl kommt der Struktur des Rücklaufs entscheidende Bedeutung zu: Nur wenn die Zusammensetzung der Teilnehmerschaft jener der Grundgesamtheit (annähernd) entspricht, sind die Ergebnisse repräsentativ und generalisierbar. Generalisieren bedeutet dann auch die Abkehr vom Leitfaden zu einem quantitativen Fragebogen. Bei Firmenbefragungen ist auch eine Online-Erhebung eine naheliegende Alternative.

Sind strukturelle Brancheninformationen nicht verfügbar, die Grundgesamtheit somit unbekannt, ist Repräsentativität zahlenmäßig nicht bezifferbar. Zunächst denkbar ist hier eine völlig willkürliche Auswahl verschiedener Firmen. Das Ergebnisbild ist dann natürlich auch ein (ziemlich) willkürliches. Eine bessere Variante stellt deshalb ein Quotenstichprobe dar. Ein möglicher Quotenplan: Jeweils zehn Experteninterviews in jedem Bundesland, die sich zu gleichen Teilen auf kleine (10 bis 49 Mitarbeitende) und mittlere Unternehmen (50 bis unter 250 Mitarbeitende) aufteilen. Mit einer Streuung über erhebungsrelevante Merkmale ist ein umfassendes, vielschichtiges Meinungsbild erzielbar. Dieses Setting erlaubt zwar kein mathematisches Generalisieren auf die unbekannt Grundgesamtheit ALLER Betriebe. Durch die breit gestreute Auswahl kommt der gewonnene Branchenüberblick einem repräsentativen Meinungsbild aber wahrscheinlich (recht) nahe: Eine Art „Repräsentativität“ ist erreichbar, bewiesen werden kann sie aber nicht.

Die größte Herausforderung für alle hier diskutierten Varianten wird wahrscheinlich das Auffinden der konkreten Expertenkontakte sein. Bei wenigen qualitativen Interviews ist

ein „Durchtelefonieren“ zur Terminvereinbarung mit der richtigen Person möglich. Bei einer quantitativen Online-Befragung empfiehlt sich in der Einladungs-E-Mail die Bitte um Weiterleitung an die richtige Person. Allerdings: Tritt jemand auf diesem Weg „im Namen der Forschung“ an zufällig ausgewählte Firmen heran, wird sich der Rücklauf ziemlich in Grenzen halten. Ist der oder die Forschende branchenbekannt, besitzt persönliches Involvement zum Thema oder stellt die Erhebung im Erstkontakt zumindest möglichst interessant dar, fördert das vermutlich die Beantwortungsquote.

Zusammenfassung möglicher Vorgehensweise(n):

- Je nach Größe und Zugänglichkeit der Grundgesamtheit Vollerhebung [43f.], Zufallsauswahl [63ff.] oder Quotenstichprobe [68ff.] nach erhebungsrelevanten Strukturmerkmalen. Ist die Grundgesamtheit unbekannt, kann Repräsentativität [55ff.] nicht argumentiert werden.
- Wenn die Auswahl keine Zufallskomponenten enthält, ist die Berechnung von Schwankungsbreiten [76ff.] eigentlich unzulässig (erfolgt aber dennoch oft als „Richtwert“ der Ergebnisgüte).
- Je nach Art der Erhebung und Offenheit der Gesprächsführung Gesprächsleitfaden [113], qualitativer Fragebogen [115] oder quantitativer Fragebogen [Beispiel 101f., im Detail 116ff.], der vereinzelt auch qualitative Elemente [123ff.] beinhalten kann.
- Die Ergebnis-Auswertung und -Darstellung erfolgt dem Setting entsprechend nach qualitativen Gesichtspunkten [134ff.] oder quantitativ [136ff. und 149ff.]. Wurden die Datensätze zufällig ausgewählt, ist bei der Ergebnisinterpretation unbedingt auf die statistischen Schwankungsbreiten zu achten [76ff.].

Mögliche Unstimmigkeiten könnten auftreten ...

... zwischen Setting und Ergebnisinterpretation: Aufgrund der schwierigen Grundgesamtheit fällt die Entscheidung auf eine qualitative Befragung von zehn Expertinnen und zehn Experten. Daraus wird ein gesamtes Branchenbild abgeleitet. → Das ist ein unzulässiges methodisches Vorgehen. Weder lassen sich qualitative Interviews geringer Menge noch willkürlich ausgewählte Stichprobenelemente verallgemeinern. Möglich sind derartige Gesamtaussagen höchstens dann, wenn die wenigen Befragten eine Totalerhebung ALLER in der Branche tätigen Personen mit Expertenstatus darstellen.

... zwischen Stichprobe und Ergebnisinterpretation: Ist die Grundgesamtheit bekannt und „erlaubt erreichbar“, kann aus ihr rasch eine Zufallsstichprobe gezogen werden. Im Beispiel werden 1.000 Expertinnen und Experten ausgewählt und online befragt. Der Rücklauf liegt über allen Erwartungen: Am Ende der Feldphase können 500 Interviews der Auswertung zugeführt werden. Damit weisen die Ergebnisse akzeptable Schwankungsbreiten auf und lassen sich treffsicher generalisieren. → Die Auswertung des Rücklaufs darf keinesfalls unkontrolliert erfolgen. Viele Antwortende bedeuten nicht unbedingt repräsentative Ergebniszahlen. Wenn die Struktur der an der Erhebung Teilnehmenden von der Grundgesamtheit stark abweicht, sind die Ergebnisse möglicherweise verzerrt: Vielleicht agieren z.B. Kleinunternehmen im Norden völlig anders als alle anderen – und gerade dieses Teilssegment hat einen viel zu geringen Anteil an den Daten. Für den Struktur-

vergleich zwischen Datenrücklauf und Grundgesamtheit müssen verfügbare und erhebungsrelevante Merkmale herangezogen werden. Entspricht die Struktur der vorliegenden Nettostichprobe nicht der Gesamtheit, werden Ergebnismessung oder Nacherhebungen unter den fehlenden Stichprobenteilen notwendig.

9.12 | Leserbefragung dieses Buchs zu diesem Buch

Die Leserinnen und Leser dieses Buchs sollen dazu befragt werden, wie ihnen Aufbau und Gestaltung gefallen.

Diese Aufgabenstellung stellt sich herausfordernd dar.

Es gibt keine bekannte Grundgesamtheit, das Buch kann ja jede und jeder irgendwo gekauft, geschenkt bekommen oder ausgeborgt haben.

In Kooperation mit einem Versandhändler an Käuferinnen und Käufer heranzutreten, mag technisch funktionieren, ist aber datenschutzrechtlich nicht möglich und liefert auch nur einen Teil der Leserschaft. Eine Vollerhebung oder Zufallsstichprobe scheiden hier ebenso aus. Auch eine Quotenstichprobe ist wenig sinnvoll – wo und wie sollen Käuferinnen und Käufer gefunden werden?

EINE Möglichkeit wären Befragungskärtchen direkt im Buch (mit Einsendeadresse) oder das Abdrucken eines Links auf ein Befragungsformular. Auch eine Befragung über die Webseite des Buchs howtodo.at oder via Social Media (im Schneeballsystem, mit darauf folgender Quotenkontrolle) könnte angedacht werden. Um hohe Diversität der Stichprobe sicherzustellen, werden aus einer Liste von 100 interessen-, alters- und regionalspezifisch für einen Themenbezug in Frage kommenden, aber dennoch möglichst unterschiedlichen Facebook-Interessengruppen 20 zufällig ausgewählt. Die Ergebnisse aus einer derartigen Erhebung sind dann zwar nicht „beweisbar repräsentativ“, können aber zumindest als Indikatoren für ein näherungsweise themenspezifisches Meinungsbild einer (facebookaffinen) Grundgesamtheit herangezogen werden.

Ein QUALITATIVES Bild liefern könnten hier inhaltliche Einzelstatements (mit Buchvorlage). Aus methodischer Sicht (in abgeschwächter Form) stellen Rezensionen eine ähnliche Erhebungsform (qualitative Inhaltsanalyse) dar.

Zusammenfassung möglicher Vorgehensweise(n):

- Willkürliche Auswahl [68] oder Quotenstichprobe [68FF.] nach zu definierenden Strukturmerkmalen, die erhebungsrelevant sind. Bei quantitativem Vorgehen zur besseren Objektivierung der Auswahl zufällige Streuung der Erhebungstage, Tageszeiten und z.B. Social Media-Bereiche [73FF.]. Repräsentativität [55FF.] kann hier nicht gemessen bzw. gewährleistet werden.
- Wenn die Auswahl keine Zufallskomponenten enthält, ist die Berechnung von Schwankungsbreiten [76FF.] eigentlich unzulässig (erfolgt aber dennoch oft als „Richtwert“ der Ergebnisgüte).
- Quantitative Erhebung [35FF.] mittels Fragebogen (Online-Formular) [116FF.] oder qualitatives Setting [34F.] mittels Gesprächsleitfaden [113] oder qualitativem Fragebogen [115].

- Die Ergebnis-Auswertung und -Darstellung [149FF.] folgen hier je nach Setting qualitativen [134FF.] oder quantitativen [136FF.] Grundsätzen.

Mögliche Unstimmigkeiten könnten auftreten ...

... **zwischen Stichprobe und Ergebnisinterpretation**: Ein eventuell verzerrtes Meinungsbild muss hier in Betracht gezogen werden: Es ist möglich, dass überdurchschnittlich häufig besonders Begeisterte und besonders Enttäuschte, eventuell auch ein paar Nicht-Lesende, an der Befragung teilnehmen – über den offenen Link vielleicht sogar mehrmals.

9.13 | Experiment zur Wirkung von Farben auf Aufmerksamkeit

Ein Möbelhersteller untersucht die Auswirkung von Farbe bzw. Holzart von Büromöbeln auf Wohlbefinden und Arbeitseffizienz. Ein Laborexperiment wird aufgesetzt.

Folgende Möblierungsvarianten sollen getestet werden: weiße, graue, rote Möbel, Möbel mit hellem, mittel- und dunkelbraunem Holzfurnier. Die Tests werden in Zusammenarbeit mit einer Studierendengruppe einer Fachhochschule durchgeführt.

Die am Experiment Teilnehmenden rekrutieren sich aus freiwilligen Studierenden der Bildungseinrichtung, an der das Experiment stattfindet. Das experimentelle Design baut auf sechs zufällig gebildeten Versuchsgruppen ohne Kontrollgruppe auf. Für das Experiment werden ein Einzelbüro, ein Doppelzimmer und ein Büro mit Vierfach-Belegung reserviert. Damit lassen sich auch eventuelle Wechselwirkungen zwischen Möblierung und Raumgröße bzw. Personenanzahl im Arbeitsraum identifizieren.

Jede Versuchsgruppe erhält per Zufall eine der sechs Möblierungsvarianten zugewiesen. Alle drei Raumgrößen werden mit dieser einen Möbelart eingerichtet. Wieder führt der Zufall Regie dabei, welche Testperson der jeweiligen Versuchsgruppe in welchem Raum (Einzel-, Doppel-, Vierer-Raum) am Experiment teilnimmt.

Das Experiment wird in jeder Versuchsgruppe (= Möbelvariante) mit unterschiedlichen Menschen fünf Mal wiederholt. Das soll die Fallbasis erhöhen und bei der Ergebnisermittlung vereinzelte individuelle Effekte in den Hintergrund drängen. Wer beim ersten Durchgang, wer beim zweiten usw. an die Reihe kommt, entscheidet ebenfalls der Zufall. Niemand nimmt doppelt (in zwei Versuchsgruppen) am Experiment teil.

Das Experiment läuft folgendermaßen ab: Die Teilnehmenden eines Durchgangs werden in die ihnen zugewiesenen Räume an ihren „Arbeitsplatz“ gebeten. In den Mehrpersonenzimmern wird die Platzwahl ebenfalls per Zufall bestimmt. Tische und Sessel sind höhenverstellbar. Die Tür zum Raum wird geschlossen. Auf jedem Arbeitsplatz liegen ein Kugelschreiber und ein auszufüllender Erhebungsbogen. ALLE am Experiment teilnehmenden Personen erhalten idente Fragen und Aufgabenstellungen. Raumklima, Wetter, Wochentag und Tageszeit sind bei allen fünf Durchgängen vergleichbar.

Die Fragen am Erhebungsbogen beschäftigen sich mit durchführungstechnischen Aspekten zum Experiment (Raumgröße, Holzfarbe) und sozialstatistischen Details. Danach erfolgt die Messung der Arbeitseffizienz. Dazu gelangt ein standardisierter psychologischer Test (in Auszügen) zur Anwendung. Im letzten Teil des Experiments wird das körperliche Wohlbefinden erfragt.

Die Auswertung vergleicht die Ergebnisse der Effizienz-Tests und die Antworten auf dem Erhebungsbogen zwischen den Versuchsgruppen. Besonderes Augenmerk wird dabei zusätzlich auf das Identifizieren eventueller Störeffekte durch soziodemografische Variablen gerichtet.

Zusammenfassung möglicher Vorgehensweise(n):

- Reine Zufallsauswahl [63F.] aus willkürlich gesammelter Teilnehmerschaft.
- Quantitativer Fragebogen [Beispiel 101F., im Detail 116FF.] bzw. Anwendung eines standardisierten Testverfahrens.
- Die Ergebnis-Auswertung und -Darstellung [149FF.] erfolgen nach quantitativen [136FF.] Grundsätzen. Da die Daten auf einer Zufallsstichprobe beruhen, ist die Anwendung statistischer Schwankungsbreiten [76FF.] und schließender Statistik möglich und sinnvoll.

Mögliche Unstimmigkeiten könnten auftreten ...

... **zwischen Setting und Ergebnisinterpretation**: Das Experiment setzt auf eine freiwillige Stichprobe (von Studierenden) in einer künstlichen Umgebung. Es „repräsentiert“ zu wenig die große Heterogenität von in Büros arbeitenden Menschen: Dazu zählen unterschiedlichste Arbeitsumgebungen, Tätigkeitsbereiche und Arbeitssituationen. → Damit sind die Ergebnisaussagen möglicherweise nur signifikant für die Versuchsbedingungen im Labor. Jedoch fehlt eine Übertragbarkeit auf reale Lebenssituationen.

9.14 | Experiment zur Wirkung von Texten auf Vertrauen

Ein Forschungsteam analysiert in einem Befragungsexperiment den Einfluss thematisierter Problemlösungskompetenz in Online-Nachrichten auf Firmen-Images.

Dazu werden fünf verschiedene Nachrichtentexte über ein fiktives Geldinstitut formuliert, das sich in wirtschaftlichen Schwierigkeiten befindet. Ein neutraler Artikel bildet die Ausgangsbasis. Er wird einer Kontrollgruppe vorgelegt. Die anderen vier Artikel thematisieren unterschiedliche Ausprägungsstufen von Problemlösungskompetenz. Jeder Artikel stellt einen Stimulus für eine von vier Versuchsgruppen dar. Die Indikatoren für die Stärke des Merkmals „Problemlösungskompetenz“ entstammen themenspezifischer wissenschaftlicher Literatur. Darauf aufbauend erfolgt die konkrete Formulierung der vier Texte.

Auch die Operationalisierung von Art und Stärke des abhängigen Merkmals Image erfolgt auf Literaturbasis. Auf Grundlage von bereits testerprobten Imagedimensionen wird eine für das gegenständliche Vorhaben leicht adaptierte Imagebatterie entwickelt. Die Imageitems messen über eine 7-stufige Zustimmungsskala (von 1 = „trifft sehr auf dieses Geldinstitut zu“ bis 7 „trifft gar nicht auf dieses Geldinstitut zu“).

Die Befragung selbst erfolgt in Form einer Online-Erhebung. Jede befragte Person erhält im Umfrageformular per Zufall EINEN der fünf Artikel eingeblendet. Dadurch „entstehen“ direkt während der Feldphase zufällig Kontrollgruppe (= Personen, die den neutralen Beitrag erhalten) und Versuchsgruppen. Die Teilnehmerinnen und Teilnehmer am Experiment werden dazu aufgefordert, ihren Artikel genau durchzulesen. Danach erhalten alle Gruppen dieselben Imageitems mit der 7-stufigen Ratingskala rotiert (= in unterschiedlicher

Reihenfolge) auf den Bildschirm. Den Abschluss des Experiments bilden soziodemografische Fragen.

Um die Erhebung auf solider Fallbasis durchführen zu können, erfolgt eine E-Mail-Adressen-Sammlung unter Studierenden. Dabei können mehrere hundert Kontaktadressen generiert werden. Auf die Definition einer Grundgesamtheit muss aus durchführungstechnischen Zwängen verzichtet werden. Dies ist aber bei derartigen Experimenten eine übliche und methodisch „tragbare“ Vorgehensweise: Es geht hier um die Analyse der Wirkung unterschiedlicher Stimuli, nicht um die Projektion von Stichprobenergebnissen auf eine Population.

Hier entscheidend ist aber eine wirklich zufällige Steuerung der Gruppenbildung (Artikelzuordnung): Dadurch werden eventuell verzerrende Effekte durch Persönlichkeitseigenschaften möglichst hintangestellt.

Die Auswertung vergleicht die in einem Gesamtscore zusammengefassten Zustimmungsbzw. Ablehnungswerte der fünf Befragtengruppen bei den Imageitems (= abhängige Variablen).

Zusammenfassung möglicher Vorgehensweise(n):

- Reine Zufallsauswahl [63F.] aus willkürlich gesammelten E-Mail-Adressen.
- Quantitativer Fragebogen [Beispiel 101F., im Detail 116FF.].
- Die Ergebnis-Auswertung und -Darstellung [149FF.] erfolgen nach quantitativen [136FF.] Grundsätzen. Da die Daten auf einer Zufallsstichprobe beruhen, ist die Anwendung statistischer Schwankungsbreiten [76FF.] und schließender Statistik möglich und sinnvoll.

Mögliche Unstimmigkeiten könnten auftreten ...

... **zwischen Fragebogen und Ergebnisinterpretation**: Die aus der Literatur abgeleiteten und für die Erhebung adaptierten Aussagen operationalisieren oder messen nicht das, was sie vorgeben. Sie sind nicht valide. → Wenn die Artikel Problemlösungskompetenz falsch oder nicht abbilden, sind keine Ergebnisaussagen zur abhängigen Dimension „Image“ möglich. Das Experiment funktioniert nicht. Dennoch gefundene (signifikante) Imagedifferenzen können nicht im Sinne des Experiments interpretiert werden. Sie sind zufällig oder in anderen, nicht definierten Ursachen begründet.

