



Leseprobe aus: Eid, Gollwitzer, Schmitt, Statistik und Forschungsmethoden, ISBN 978-3-621-28201-7
© 2015 Beltz Verlag, Weinheim Basel
<http://www.beltz.de/de/nc/verlagsgruppe-beltz/gesamtprogramm.html?isbn=978-3-621-28201-7>

2 Struktur und Ablauf wissenschaftlicher Untersuchungen

Was Sie in diesem Kapitel lernen

- ▶ Was ist eine wissenschaftliche Hypothese, und wie testet man sie?
- ▶ Hat Sigmund Freud jemals einen Abwehrmechanismus gesehen?
- ▶ Wie lässt sich wissenschaftlich ermitteln, ob eine Person tabuisierte sexuelle Wünsche verdrängt hat?
- ▶ Wie könnte man wissenschaftlich herausfinden, welche Bedeutung das Aufschlitzen von Sitzen in öffentlichen Verkehrsmitteln hat?

Um die Bedeutung von Forschungsmethoden für die Psychologie besser zu verstehen, ist es hilfreich, sie in den Gesamtprozess von Wissenschaft einzuordnen. Wir wollen diese Einordnung hier in zweierlei Hinsicht vornehmen: Zunächst werden wir uns die verschiedenen Ebenen, auf denen wissenschaftliche Aussagen getroffen werden, und die Beziehungen, die zwischen diesen Ebenen bestehen, etwas genauer ansehen. In einem zweiten Schritt werden wir den Ablauf einer typischen wissenschaftlichen Untersuchung nachzeichnen und dabei herausarbeiten, an welchen Stellen welche Methoden in welcher Funktion angewendet werden.

2.1 Hypothesen, Ebenen wissenschaftlicher Aussagen und die Überbrückungsproblematik

Jede Erfahrungswissenschaft trifft, bearbeitet und testet Aussagen auf unterschiedlichen Ebenen. Auf der theoretisch-konzeptionellen Ebene werden Aussagen über Objekte und Ereignisse (in theoretischen Begriffen) sowie die Beziehungen zwischen diesen Objekten und Ereignissen (in Form theoretischer Zusammenhänge) getroffen. Eine theoretische Aussage könnte z. B. lauten: »Wer intelligent ist, der ist auch kreativ.« Intelligenz und Kreativität sind theoretische Begriffe. Die Aussage stellt eine Behauptung dar, deren Wahrheitsgehalt empirisch geprüft werden kann.

2.1.1 Prüfbar und nicht-prüfbar Aussagen

Der Begriff »Empirie« stammt vom griechischen Substantiv »empeiria« und bedeutet Erfahrung im Sinne von sinnlicher Wahrnehmung. Nicht alle Aussagen sind empirisch prüfbar. Manche Aussagen sind praktisch nicht prüfbar. Die Aussage »In der Hölle ist es heißer als 220 Grad« ist nicht prüfbar, da es sich bei der Hölle um einen nicht ohne Weiteres zugänglichen Ort (viele Menschen würden sogar sagen: gar keinen real existierenden Ort) handelt. Mit einem Thermometer in der Hölle Forschung zu betreiben ist also nicht möglich. Etwas technischer gesprochen: Die thermischen Zustände in der Hölle sind nicht erfahrbare. Erfahrbare ist jedoch in den empirischen Wissenschaften eine notwendige Grundvoraussetzung für die Überprüfung theoretischer Aussagen.

Ebenfalls nicht empirisch prüfbar sind Aussagen, die per Definition richtig sein *müssen*. Unverheiratete Männer bezeichnet man als Junggesellen. Empirisch zu prüfen, ob Junggesellen verheiratet sind, ist nicht möglich: Wenn die Definition richtig angewendet wird, kann es keine verheirateten Junggesellen geben. Empirisch nicht prüfbar sind weiterhin Aussagen, die nicht falsch sein können. Dazu gehören Möglichkeitssätze wie »Frauen können schwanger werden«.

Falsifizierbarkeit. Die Feststellung, dass manche Aussagen nicht prüfbar sind, weil sie immer richtig sind, lässt schon erahnen, worauf es bei der Prüfung einer theoretischen Aussage ankommt: auf ihre Falsifizierbarkeit. Eine Aussage muss prinzipiell falsch sein *können*, damit sie empirisch prüfbar ist. Ergibt die Prüfung, dass sie tatsächlich falsch ist, dann hat man ein eindeutiges Ergebnis. Solange jedoch eine Aussage nicht falsifiziert werden kann, muss sie als vorläufig gültig betrachtet werden. Auf die Einschränkung »vorläufig« darf nur dann verzichtet werden, wenn alle Möglichkeiten zur Falsifizierung der Aussage ausgeschöpft sind. Manche Aussagen sind praktisch nicht falsifizierbar. Nehmen wir die Aussage: »Es gibt Menschen, die dreimal am Tag warm essen und trotzdem kein Gramm zunehmen.« Man könnte diese Aussage nur dann falsifizieren, wenn

man nachweisen könnte, dass kein Mensch auf der Welt bei dreimaligem Essen am Tag auch nur ein Gramm zunimmt. Eine solche Form der Falsifizierung ist praktisch unmöglich, weil man dazu alle Menschen, die gewöhnlich auf unserer Welt leben, untersuchen müsste.

2.1.2 Wissenschaftliche Hypothesen

Ein zentrales wissenschaftstheoretisches Konzept ist die Hypothese. Der Begriff »Hypothese« stammt ebenfalls aus dem Griechischen (*hypóthesis*) und bedeutet Unterstellung, Voraussetzung oder Grundlage. Hypothesen sind Aussagen, die empirisch testbar und somit falsifizierbar sein müssen. Darüber hinaus kommt einer testbaren Aussage nur dann der Rang einer wissenschaftlichen Hypothese zu, wenn sie eine gewisse Allgemeingültigkeit beansprucht, also über den Einzelfall hinausweist. Die Aussage »Peter wird morgen Kopfschmerzen haben, wenn er weiter so viel Alkohol trinkt« ist also noch keine wissenschaftliche Hypothese. Allerdings verbirgt sich hinter dieser Aussage eine (vermutete) allgemeine Gesetzmäßigkeit, nämlich die, dass übermäßiger Alkoholenuss am Abend im Allgemeinen zu Kopfweh am Morgen führt. Bei dieser Aussage handelt es sich dann um eine wissenschaftliche Hypothese.

Darüber hinaus muss eine wissenschaftliche Hypothese begründet sein. Die Aussage »Wer sich die Haare rot färbt, isst mehr Fisch« wäre zwar testbar und hat eine ausreichende Allgemeingültigkeit, aber man würde dieser Aussage nicht den Rang einer wissenschaftlichen Hypothese zusprechen, solange sie nicht begründet ist.

Definition

Bei einer Aussage handelt es sich um eine wissenschaftliche Hypothese, wenn sie prinzipiell der sinnlichen Erfahrung zugänglich ist, prinzipiell widerlegbar ist, eine gewisse Allgemeingültigkeit beansprucht und theoretisch begründet ist.

Empirische Ebene

Um eine theoretische Aussage bzw. eine wissenschaftliche Hypothese empirisch zu testen, muss man die theoretischen Konzepte und die theoretischen Relationen zwischen diesen Konzepten in konkrete Aussagen überführen, deren Konzepte und Relationen empi-

risch erfahrbar sind. Dies erfordert einen nicht immer einfachen Übersetzungsprozess, den wir an einem prominenten Beispiel veranschaulichen wollen: dem Konzept des Abwehrmechanismus, einem zentralen Bestandteil der psychoanalytischen Theorie des österreichischen Arztes und Tiefenpsychologen Sigmund Freud (1856–1939).

Theoretische Hypothesen

Abwehrmechanismus ist ein theoretischer Begriff. Niemand hat jemals einen Abwehrmechanismus gesehen, auch Freud nicht. Das Konzept des Abwehrmechanismus ist ein gedankliches Gebilde, eine Schöpfung des Freud'schen Geistes, ein hypothetisches Konstrukt. Freud hat dieses Konstrukt geschaffen, um Beobachtungen, die er in seiner Praxis an Klienten gemacht hat, zu deuten, ihnen einen psychologischen Sinn zu geben.

Freud stellte in seiner praktischen Arbeit häufig fest, dass seine Klienten bestimmte Episoden aus ihrem Leben vergessen hatten und sich erst im Prozess der Psychoanalyse wieder daran erinnerten. In diesen Episoden kamen meistens Handlungen, Vorstellungen, Wünsche oder Fantasien mit sexuellen Inhalten vor. Freud nahm an, dass sexuelle Themen, insbesondere bestimmte sexuelle Wunschvorstellungen (z. B. die eines Knaben, mit der eigenen Mutter erotischen Kontakt zu haben), der Person Angst machen, weil sie tabuisiert sind und sanktioniert werden. So habe etwa der Knabe, der die eigene Mutter begehrt, Angst, vom Vater kastriert zu werden. Abwehrmechanismen dienen nun nach Freuds theoretischen Vorstellungen dazu, solche Ängste zu bewältigen. Eine Möglichkeit besteht darin, die sexuellen Wunschvorstellungen zu verdrängen. Verdrängung ist einer von mehreren Abwehrmechanismen und selbst wiederum ein theoretischer Begriff. Die Verdrängung leistet eine Verschiebung der tabuisierten Wunschvorstellung ins Unbewusste. Auch beim Unbewussten handelt es sich um einen theoretischen Begriff. War die Verdrängung erfolgreich, wird die tabuisierte Wunschvorstellung vergessen und kann erst wieder mit den Mitteln der psychoanalytischen Behandlung (z. B. der freien Assoziation oder der Traumdeutung) der bewussten Erinnerung zugänglich gemacht werden.

Aus diesem Teil der Theorie Freuds lassen sich zahlreiche theoretische Hypothesen ableiten. So sollten tabuisierte Wunschvorstellungen umso wahrscheinlicher verdrängt werden, je stärker sie sind. Daraus folgt, dass der Aufwand an psychoanalytischer Arbeit, der erfor-

derlich ist, um eine tabuisierte Wunschvorstellung in Erinnerung zu rufen, mit der Stärke dieser Wunschvorstellung steigt. Vermittelt werden diese Zusammenhänge theoretisch durch die Angst vor Sanktionen, die bei starken Tabuwünschen größer sein sollte als bei schwachen.

2.1.3 Überbrückungsprobleme

Theorie und Empirie

Um diese theoretischen Hypothesen empirisch prüfen zu können, müssen die Bestandteile der Hypothesen konkretisiert werden. Die Hypothese kann also nur bei konkreten Personen, mit konkreten Tabuwünschen, konkreten Formen der psychoanalytischen Behandlung und einer konkreten Form des Aufwands psychoanalytischer Arbeit angewendet werden. Diese Konkretisierungen sind aber kein Bestandteil der Theorie und lassen sich auch nicht unmittelbar aus den theoretischen Sätzen ableiten. Vielmehr handelt es sich um einen Zuordnungsprozess, in den viele Überlegungen und Entscheidungen einfließen, die außerhalb der Theorie stehen. So sagt die Theorie nichts darüber aus, bei welchen Personen welche Wünsche tabu sind, welche Personen unter welchen Umständen von welchem Mechanismus zur Abwehr ihrer Angst Gebrauch machen und welches Element der psychoanalytischen Theorie bei welchem Tabuthema, bei welchem Klienten, bei welchem Abwehrmechanismus wie effektiv ist.

Eine Übertragung der theoretischen Hypothese in konkrete empirische Hypothesen ist also mit Unsicherheiten behaftet. Die Theorie lässt ihrer Anwendung auf die Wirklichkeit Spielraum und zwingt den Anwender ebenso wie den Forscher, der die Theorie überprüfen möchte, zu Festlegungen, die bis zu einem gewissen Grade beliebig sind. Dies erkennt man u. a. daran, dass unterschiedliche Forscher die gleiche Theorie unterschiedlich anwenden und unterschiedliche empirische Beobachtungen für geeignet halten, die Theorie zu prüfen.

Nehmen wir einmal an, ein Forscher habe sich zur empirischen Prüfung der fraglichen Hypothese entschlossen und die unvermeidlichen Festlegungen getroffen. Beispielsweise habe er sich entschlossen, die Theorie am Beispiel des tabuisierten Wunsches von Knaben nach einer erotischen Beziehung zur eigenen Mutter zu prüfen und die Stärke der Verdrängung über den psychotherapeutischen Aufwand zu bestimmen,

der erforderlich ist, um die Erinnerung an das erotische Begehren wieder herzustellen. Nach dieser Festlegung muss sich der Forscher auf die Suche nach Therapeuten machen, die bereit sind, ihm Klienten zu vermitteln, bei denen sich die Hypothese untersuchen lässt. Auch für diese Entscheidung macht die Theorie keine konkreten Vorgaben.

! Das **erste Überbrückungsproblem** besteht darin, die Bestandteile einer theoretischen, wissenschaftlichen Aussage in konkrete, empirisch erfahrbare Aussagen zu überführen. Darüber, wie solche Konkretisierungen zu treffen sind, machen Theorien im Allgemeinen keine Aussage.

Operationalisierung und Messung

Kommen wir zum zweiten Überbrückungsproblem, zu der Übersetzung empirischer Aussagen in die Sprache der Zahlen. Wie kann und wie sollte der Forscher, wenn er genügend Klienten gefunden hat, die zur Teilnahme an der Untersuchung bereit sind, die relevanten Größen, die in der Hypothese vorkommen, registrieren, die erhobenen Informationen in die Symbolsprache der Zahlen übersetzen und diese Zahlen auf eine Weise auswerten, die eine bestmögliche Prüfung der Hypothese gewährleistet?

Operationalisierung. Auf diese Frage gibt es viele Antworten, und es ist schwer zu entscheiden, welche davon die beste ist. Beispielsweise könnte der Forscher so vorgehen, dass er die Mütter der Klienten dazu befragt, wie oft ihr Sohn in einem bestimmten Moment Kontakt zu ihr gesucht hat, den man als erotische Annäherung bezeichnen könnte. Blenden wir die offensichtlichen Schwierigkeiten, die mit dieser Methode verbunden sind, einmal aus und nehmen an, die Mütter der Klienten würden die Frage beantworten. Dann könnte der Forscher die Häufigkeit von Annäherungsversuchen als Maß für die Stärke des tabuisierten erotischen Begehrens nehmen. Eine solche Art der Übersetzung wird in der Methodensprache häufig als Operationalisierung (Messbarmachung) bezeichnet. Später werden wir den Begriff des Messens präzise definieren und ausführlich erläutern (Abschn. 5.2).

Im gegenwärtigen Beispiel stellt die Menge der natürlichen Zahlen die Symbole bereit, in denen wir den ersten Teil der zu prüfenden Behauptung formulieren

könnten. In ähnlicher Weise könnte der Forscher sich entscheiden, den psychotherapeutischen Aufwand durch die Zahl der therapeutischen Sitzungen anzugeben, die laut den Aufzeichnungen der Psychoanalytiker erforderlich waren, um den verdrängten Tabubruch in Erinnerung zu rufen. Auch die mit dieser Entscheidung verbundenen Probleme lassen wir im Moment beiseite und halten fest, dass sich auch der zweite Teil der theoretisch erwarteten empirischen Verhältnisse in der Sprache der natürlichen Zahlen ausdrücken ließe.

! Das **zweite Überbrückungsproblem** besteht darin, diejenigen Bestandteile, über deren Relation die Hypothese eine Aussage macht, zu quantifizieren, d. h. in messbare Größen (die wiederum mithilfe von Zahlen darstellbar sind) zu übertragen. Darüber, wie solche Operationalisierungen vorzunehmen sind, machen Theorien im Allgemeinen keine Aussage.

Datenerhebung und Datenauswertung

Nehmen wir schließlich an, der Forscher habe die beiden fraglichen Größen auf die beschriebene Weise bei einer Gruppe von Klienten ermittelt. Er verfügt dann über zwei Zahlenreihen, die mit einer der Hypothese angemessenen Methode ausgewertet werden müssen.

Was bedeuten in diesem Zusammenhang »auswerten« und »angemessen«? Auswerten heißt, die Zahlen in einer Weise zu ordnen, dass sie eine Aussage ergeben. Angemessen heißt, dass die Form dieser Aussage jener der empirischen Aussage möglichst genau entspricht, also der Aussage, dass die Häufigkeit von Annäherungsversuchen als Maß für die Stärke des tabuisierten erotischen Begehrens mit der Zahl der therapeutischen Sitzungen zusammenhängt, die laut den Aufzeichnungen der Psychoanalytiker erforderlich waren, um den verdrängten Tabubruch in Erinnerung zu rufen. Wie wir später erfahren werden, könnte die Korrelationsanalyse diesen Zweck erfüllen. Auf der Basis des Ergebnisses dieser Analyse würde der Forscher beurteilen, ob bzw. wie genau die theoretischen Vorhersagen eingetroffen sind.

Die Qualität der Schlussfolgerung, die der Forscher aus seinen Daten zur Beurteilung der Theorie zieht, hängt davon ab, wie gut er die Theorie in empirische Hypothesen übersetzt hat, wie gut seine empirischen Beobachtungen waren, wie gut er seine empirischen

Beobachtungen in die Zahlensprache übersetzt hat und wie gut sich die gewählte Auswertungsmethode für die verfügbaren Daten und die Fragestellung eignet. Der Forscher – und dies gilt für den Praktiker in vergleichbarer Weise – muss also im Forschungsprozess zahlreiche Entscheidungen fällen, deren Güte sich nicht in jedem Fall mit letzter Sicherheit beurteilen lässt. Dies ist einer der vielen Gründe, weshalb Theorien durch empirische Untersuchungen nie definitiv und abschließend als richtig oder falsch beurteilt werden können. Wenn z. B. die Daten mit einer unpassenden oder mangelhaften Methode erhoben oder ausgewertet wurden, sagen die Ergebnisse nichts über die Theorie aus, die geprüft werden sollte. Deshalb ist es wichtig, die besten der verfügbaren Methoden für empirische Untersuchungen zu finden und den Untersuchungsprozess so transparent wie möglich zu gestalten, damit andere Wissenschaftler oder Praktiker die Entstehung der Ergebnisse nachvollziehen, sich ein eigenes Urteil über die Qualität der Untersuchung bilden und die Fragestellung gegebenenfalls mit besseren Methoden erneut untersuchen können.

2.2 Schritte im Forschungsprozess

Das Beispiel, mit dem wir im letzten Abschnitt die Ebenen wissenschaftlicher Aussagen und die Schwierigkeiten, diese Ebenen ineinander zu überführen, illustriert haben, ließ bereits in groben Zügen den Ablauf einer wissenschaftlichen Untersuchung erkennen. Betrachten wir nun diesen Ablauf etwas genauer.

2.2.1 Entstehung eines Erkenntnisinteresses

Der Forschungsprozess beginnt in der Regel mit einem Erkenntnisinteresse. Dessen Quellen können vielfältig sein:

- ▶ Man stößt auf Ungereimtheiten in den Befunden verschiedener Untersuchungen und möchte sie klären.
- ▶ Man glaubt einen Befund nicht, den man liest, und möchte ihn selbst nachprüfen.
- ▶ Man erhält die Anfrage eines Praktikers, der mit einem Problem konfrontiert ist, für das er keine Lösung kennt.
- ▶ Man erhält von einem Auftraggeber den gezielten Auftrag, eine Fragestellung zu klären.

- ▶ Man macht im Alltag eine Beobachtung, die man sich nicht erklären kann und für die man auch in der Literatur keine Erklärung findet.
- ▶ Man stößt in der Literatur auf viele unterschiedliche Erklärungen für ein Phänomen, das man gerne verstehen möchte.

Die eigene Forschung könnte dann das Ziel verfolgen herauszufinden, welche Erklärung sich am besten bewährt oder unter welchen Randbedingungen welche Erklärung zutrifft.

Nehmen wir einmal an, eine Forscherin, die normalerweise keine öffentlichen Verkehrsmittel benutzt, würde ausnahmsweise an mehreren Tagen mit der Straßenbahn fahren und dabei würde ihr auffallen, dass viele Sitzbezüge aufgeschlitzt sind. Nachdem sich ihre erste Empörung gelegt hat, beginnt die Forscherin neugierig zu werden und sich nach den Ursachen zu fragen. Sie beschließt, der Frage wissenschaftlich nachzugehen.

2.2.2 Sammlung verfügbaren Wissens

Die meisten Wissenschaftlerinnen und Wissenschaftler werden in einer solchen Situation damit beginnen nachzudenken, ob sie über das erklärungsbedürftige Phänomen vielleicht schon etwas wissen. Falls ihnen keine Theorie oder Untersuchung zum Thema einfällt, werden sie vielleicht das Gespräch mit Kolleginnen oder Kollegen suchen, denen sie zutrauen, etwas von der Sache zu verstehen.

Gleichzeitig wird man sich als Wissenschaftler in der Literatur auf die Suche nach Arbeiten machen, die sich mit dem Phänomen oder verwandten Phänomenen befasst haben. Geeignet für diesen Zweck sind Literaturdatenbanken, Lehr- und Handbücher sowie Zeitschriften, die Überblicksarbeiten (sog. Review-Artikel) publizieren (z. B. die Zeitschrift »Annual Review of Psychology«). Mithilfe von passenden Suchbegriffen – im gegenwärtigen Beispiel etwa »Vandalismus«, »mutwillige Zerstörung« oder »Zerstörung öffentlichen Eigentums« – wird man versuchen, möglichst aktuelle Arbeiten zum Thema zu finden. Wird man fündig, beginnt man diese Arbeiten zu lesen und prüft dabei, ob sie das eigene Erkenntnisinteresse hinreichend befriedigen können oder nicht. Meistens wird man in den gelesenen Artikeln Hinweise auf weitere Arbeiten finden, die man mit der ersten Suchstrategie nicht entdeckt hat.

Auch diese Arbeiten wird man lesen und nach brauchbaren Informationen für die eigene Fragestellung durchforsten. Wenn das Erkenntnisinteresse im Zuge dieser Lektüre befriedigt werden kann, hat sich eine eigene Untersuchung erübrigt. Häufig jedoch bleiben Fragen offen oder es entstehen neue. In diesem Prozess kann es also bereits zu einer Veränderung oder Präzisierung des eigenen Erkenntnisinteresses kommen.

Eine weitere Strategie in dieser frühen Erkundungsphase besteht darin, zu überlegen und systematisch anhand der verfügbaren Literatur zu prüfen, ob sich Erkenntnisse aus anderen Gebieten der Psychologie, die vordergründig mit dem interessierenden Phänomen nichts zu tun zu haben scheinen, übertragen lassen. Möglicherweise gehört das Aufschlitzen von Sitzen in öffentlichen Verkehrsmitteln zu einer Kategorie, die viele psychologisch gleichwertige Verhaltensweisen umfasst. Und möglicherweise gibt es zu anderen Verhaltensweisen aus dieser Kategorie bereits brauchbare Erkenntnisse, die sich auf die eigene Fragestellung anwenden lassen. Probehalter könnte man das Aufschlitzen von Sitzen als eine von vielen Formen des Auslebens von Aggressionen oder als eine von vielen normabweichenden Verhaltensweisen interpretieren. Dann würde man den Erkundungsprozess auf diese allgemeineren Phänomene (Verhaltensklassen) ausdehnen und erneut in die Literatursuche eintreten, diesmal jedoch mit allgemeineren Suchbegriffen wie »Aggression« oder »normabweichendes Verhalten«.

2.2.3 Entwicklung einer Fragestellung oder Hypothese

In vielen Fällen wird man bei einer solchen erweiterten Suche nach bereits verfügbaren Erkenntnissen fündig. In manchen Fällen wird man sie für so gut übertragbar halten, dass sich eine eigene Untersuchung erübrigt. Andernfalls lässt man sich durch die gefundenen Arbeiten dazu anregen, vorhandene Theorien auf die eigene Fragestellung zu übertragen. Diese Übertragung ist jedoch mit einem Risiko behaftet: Es bleibt die Ungewissheit, ob sich die Theorie, die sich in anderen Anwendungen bewährt hat, auch für die eigene Fragestellung eignet. Eine eigene Untersuchung könnte dann das Ziel haben, diese Ungewissheit zu beseitigen.

Für den Fall, dass sowohl die spezifische als auch die erweiterte Literatursuche ergebnislos bleibt, müssen

neue Erklärungen entwickelt werden. Meistens wird man in einem solchen Fall jedoch nicht gleich eine ausgearbeitete Theorie anstreben, sondern zunächst die Beobachtung des erklärungsbedürftigen Phänomens ausdehnen und systematischer vornehmen. Dies geht zwar nicht ganz ohne theoretische Vorannahmen; die Untersuchung verfolgt dennoch eher die Klärung einer Fragestellung als die Prüfung einer ausgearbeiteten Theorie.

In unserem Beispiel könnten sinnvolle Fragestellungen etwa lauten:

- ▶ Zeigen Personen, die Sitze aufschlitzen, noch andere Verhaltensweisen, die man in die gleiche psychologische Kategorie einordnen kann? Beschädigen sie z. B. auch Straßenlampen und öffentliche Toiletten?
- ▶ Richten sich die Verhaltensweisen nur gegen öffentlichen oder auch gegen privaten Besitz? Tendieren die betreffenden Personen z. B. auch dazu, Autoantennen und Scheibenwischer an privaten Pkws abzubrechen?
- ▶ Werden von den Personen auch andere normwidrige Verhaltensweisen gezeigt, die nichts mit Vandalismus zu tun haben? Begehen sie z. B. auch Ladendiebstähle, erpressen sie ihre Mitschüler oder Arbeitskollegen, hinterziehen sie Steuern, begehen sie Versicherungsbetrug?

Hinter solchen Fragestellungen stehen in aller Regel schon Vermutungen, die mehr oder weniger konkret und präzise sein können. Beispielsweise könnte hinter der dritten Fragestellung die Annahme stehen, dass es eine interindividuell unterschiedliche und über viele Verhaltensbereiche generalisierte Bereitschaft gibt, soziale Normen zu übertreten.

In diesem Falle hätte die Fragestellung bereits den Charakter einer konkreten Hypothese: Das Aufschlitzen von Sitzen ist eines von vielen Anzeichen einer Disposition (Neigung) zu abweichendem Verhalten. Dies ist jedoch nur eine von vielen denkbaren Erklärungen für das beobachtete Phänomen. Weitere Hypothesen könnten lauten:

- ▶ Die Täter wissen vielleicht nicht, dass das Aufschlitzen von Sitzen unerwünscht und verboten ist.
- ▶ Das Aufschlitzen von Sitzen könnte eine Mutprobe sein und dazu dienen, Freunden zu imponieren.
- ▶ Das Aufschlitzen von Sitzen könnte ein Zeichen von Neugier sein. Vielleicht haben die Täter ein brennendes Interesse herauszufinden, wie die Sitze aufgebaut

sind, wie robust sie sind oder wie leistungsfähig das neu erworbene Taschenmesser ist.

- ▶ Das Aufschlitzen von Sitzen könnte eine symbolische Botschaft an andere Fahrgäste sein, dass man sich der Gesellschaft nicht zugehörig fühlt und ihre Spielregeln nicht akzeptiert.

Viele weitere Erklärungen sind denkbar. Wichtig ist es, bei der Formulierung von Hypothesen im Auge zu behalten, dass sie empirisch prüfbar bzw. widerlegbar sind, eine gewisse Allgemeingültigkeit aufweisen und theoretisch begründet sind. Die theoretische Begründung ergibt sich meist aus der Literaturrecherche. Um Hypothesen einer empirischen Prüfung unterziehen zu können, ist es ferner von Vorteil, alle theoretischen Konzepte, deren man sich bedient, so präzise wie möglich zu definieren. Im Falle der Hypothese, das Aufschlitzen von Sitzen diene dazu, Freunden zu imponieren, müsste definiert werden, was unter »imponieren« zu verstehen ist. Im Falle der Hypothese, das Aufschlitzen von Sitzen befriedige Neugier, müsste definiert werden, was »Neugier« bedeutet, etc. Je präziser die theoretischen Konzepte, die Bestandteil einer Hypothese sind, definiert sind, desto leichter fällt die Hypothesenprüfung.

2.2.4 Planung einer Untersuchung

Um solche Fragestellungen bzw. Hypothesen zu klären, bedarf es nach dem Selbstverständnis unserer wissenschaftlichen Disziplin einer empirischen Untersuchung. Zur weiteren Illustration des Forschungsprozesses wählen wir die Dispositionshypothese aus, die besagt, dass das Aufschlitzen von Sitzen eines von vielen Anzeichen einer generalisierten Bereitschaft zur Übertretung von sozialen Normen ist. Die Konzentration auf diese Hypothese macht es zunächst erforderlich, andere Formen normabweichenden Verhaltens zu bestimmen. Einige Beispiele wurden oben gegeben (Versicherungsbetrug etc.).

Auswahl einer Erhebungsmethode

Anschließend muss überlegt und entschieden werden, wie man über diese Verhaltensweisen Informationen gewinnen möchte. Beispielsweise könnte man sich für eine Verhaltensbeobachtung oder für eine Befragung entscheiden. Weitere Möglichkeiten werden wir später kennenlernen (Kap. 3). Würde man sich für eine

Befragung entscheiden, müsste man einen Fragebogen entwickeln, der eine Vielzahl normabweichender Verhaltensweisen enthält. Man könnte zu jeder Verhaltensweise fragen, ob sie schon einmal gezeigt wurde, wann das letzte Mal und wie oft in einem definierten Zeitraum. Da normabweichendes Verhalten von den meisten Menschen aus Angst vor Strafe oder Ablehnung nur ungern zugegeben wird, müsste man mit geeigneten Maßnahmen sicherstellen, dass die befragten Personen ehrlich antworten. Dies könnte man etwa durch anonyme Befragungen erreichen, wie sie in der Kriminologie zur Ermittlung von Dunkelziffern durchgeführt werden. Alternativ dazu könnte man die Aussagen der Personen überprüfen, indem man Bezugspersonen wie Freunde, Verwandte und Lehrer über das gleiche Verhalten der Person befragt und die Angaben der Zielperson mit jenen dieser Fremdbeurteiler vergleicht.

Festlegung der Population und Auswahl einer Stichprobe

Als Nächstes muss für die Erhebung eine Stichprobe ausgewählt und deren Größe festgelegt werden. Wie wir später lernen werden, ist es dabei wichtig, dass man zunächst die Grundgesamtheit (Population) derjenigen Personen definiert, für die die zu prüfende Hypothese gelten soll. In unserem Beispiel könnte die Grundgesamtheit aus Jugendlichen und jungen Erwachsenen bestehen. Aus dieser Grundgesamtheit muss dann eine repräsentative Stichprobe gezogen werden. Die Stichprobengröße ergibt sich aus dem Genauigkeitsanspruch der Untersuchung. Wenn die Untersuchung eine erste Erkundung sein soll und v. a. den Zweck verfolgt, weiterführende Ideen zu generieren, genügt eine kleine Zahl von Personen. Wenn hingegen eine Theorie getestet werden soll oder aus den Befunden weitreichende Schlussfolgerungen gezogen werden sollen, von denen die Grundgesamtheit betroffen ist, wird die Stichprobe größer sein müssen. Warum das so ist und wie man bei der Planung der Stichprobengröße vorgeht, werden wir später erfahren (Kap. 8).

Probleme bei der Versuchsdurchführung

Nachdem man sich für eine bestimmte Form der empirischen Prüfung der Hypothese entschieden hat, sollte man sich als Nächstes die Frage stellen, mit welchen Problemen bei der Versuchsdurchführung potenziell zu rechnen ist.

Mangelnde Validität. Eine Schwierigkeit haben wir bereits kennengelernt: Fragt man Menschen nach ihrer Neigung zu normabweichendem Verhalten, besteht die Gefahr, dass man keine ehrlichen Antworten erhält; es könnte also sein, dass zwei Personen, die auf die Frage, wie oft sie im letzten Jahr Ladendiebstahl begangen haben, »noch nie« bzw. »drei Mal« antworten, sich nicht wirklich in ihrer Normbruchneigung unterscheiden, sondern lediglich in ihrer Ehrlichkeit. Das wäre misslich, denn man hätte nicht das gemessen, was man messen wollte: Statt Unterschieden in der Normbruchneigung hätte man Unterschiede in der Ehrlichkeit erfasst. Wir werden später sehen, dass das Ausmaß, in dem eine empirische Messung tatsächlich das erfasst, was sie erfassen soll, eine Eigenschaft der Messung ist und als »Validität« bezeichnet wird (Abschn. 3.4). Wir werden auch sehen, wie man die Validität einer Messung quantifizieren kann und mit welchen Strategien man die Validität der Messung erhöhen kann.

Systematisch fehlende Werte. Eine weitere Schwierigkeit könnte darin bestehen, dass Personen, die man in die Stichprobe gezogen hat, sich weigern, die gestellten Fragen zu beantworten. Dann gäbe es fehlende Werte in den Daten. Solange solche fehlenden Werte unsystematisch über die Personen hinweg verteilt sind und es mehr oder weniger Zufall ist, wer die Auskunft verweigert und wer nicht, gibt es keine großen Probleme. Schwierig wird es hingegen, wenn nur bestimmte Personen die Auskunft verweigern, z. B. diejenigen, die in sehr starkem Maße zu normabweichendem Verhalten neigen. Wenn also Werte aus einem bestimmten Spektrum systematisch fehlen, kann das die Aussagekraft der Daten erheblich mindern.

Ethische Unbedenklichkeit

Eine zentrale Frage, die man sich bei der Planung der eigenen Untersuchung stellen sollte, ist die der ethischen Unbedenklichkeit. Wissenschaftlerinnen und Wissenschaftler haben ein Erkenntnisinteresse, das im Idealfall der Gesellschaft dient, Fragen beantwortet und Probleme lösen hilft. Aber manchmal steht dieses Erkenntnisinteresse im Konflikt mit Prinzipien der ethischen Verantwortlichkeit gegenüber den Versuchspersonen. Ethisch bedenklich sind Untersuchungen, wenn sie die Menschenwürde verletzen oder mit potenziellen Gefahren für Leib, Leben und Wohlergehen verbunden sind. Die Frage, wie Menschen Traumata und schweren

psychischen Stress bewältigen, mag wissenschaftlich und gesellschaftlich hoch relevant sein, aber sie berührt einen sensiblen Bereich. So verbietet es sich, Traumata gezielt experimentell auszulösen, um zu untersuchen, wie deren kognitive Verarbeitung funktioniert. Auch die bloße Befragung zu zurückliegenden traumatischen Erfahrungen kann ethisch bedenklich sein, wenn die Gefahr besteht, dass die Befragungssituation von der befragten Person als belastend erlebt wird. Ethisch bedenklich wäre in diesem Fall andererseits der Verzicht auf eine Untersuchung, deren Ergebnisse den betroffenen Personen helfen könnten, ihre traumatischen Erfahrungen besser zu verarbeiten.

Die Deutsche Gesellschaft für Psychologie (DGPs) und der Berufsverband Deutscher Psychologinnen und Psychologen (BDP) haben gemeinsam ethische Richtlinien herausgegeben, die nicht nur die Unbedenklichkeit wissenschaftlicher Fragestellungen und empirischer Herangehensweisen betreffen, sondern auch allgemeine Fragen der beruflichen Praxis von Psychologinnen und Psychologen. Der aktuelle Text ist im Internet abrufbar. ([📄](#) Einen Link darauf finden Sie in unseren Online-Materialien.)

2.2.5 Durchführung der Untersuchung

Zunächst muss das Untersuchungsmaterial erstellt werden, z.B. der Fragebogen, mit dem die Häufigkeit normabweichenden Verhaltens ermittelt werden soll. Die Konstruktion von Fragebögen ist eine Kunst für sich, die wir hier nicht im Detail behandeln können. Auf einige Aspekte, die es dabei zu beachten und entscheiden gilt, werden wir später genauer eingehen (Abschn. 3.3.3). Außerdem gibt es Bücher, die sich ausschließlich mit der Konstruktion von Fragebögen befassen (z.B. Moosbrugger & Kelava, 2012; Mummeny & Grau, 2014).

Als Nächstes wird die Stichprobe rekrutiert. In unserem Beispiel könnte man etwa versuchen, über Schulen an die jüngeren Untersuchungsteilnehmer zu gelangen. Man könnte aber auch Personen auf der Straße oder in öffentlichen Verkehrsmitteln ansprechen und um ihre Teilnahme bitten. Weiterhin könnte man sich von der Einwohnerbehörde eine Stichprobe von Personen aus dem Melderegister ziehen lassen. Schließlich könnte mithilfe von Anzeigen in der Tagespresse zur Teilnahme an der Untersuchung aufgerufen werden.

Welche dieser Strategien man verfolgt und wie man sie konkret anwendet, hängt von der Fragestellung, der Definition der Grundgesamtheit sowie von Annahmen über Gründe der Teilnahmebereitschaft und befürchteten Störfaktoren ab. Rechnet man z. B. mit regionalen Unterschieden im fraglichen Verhalten, muss man die Region bei der Stichprobenziehung systematisch berücksichtigen.

Sobald die Stichprobe gezogen ist, kann mit der Datenerhebung begonnen werden. In unserem Beispiel bedeutet dies, dass man den Fragebogen in Schulklassen austeilt, jenen Personen, die man auf der Straße oder in öffentlichen Verkehrsmitteln angesprochen hat, den Fragebogen nebst einem frankierten Rücksendeumschlag aushändigt oder jenen Personen, deren Anschrift man von der Einwohnerbehörde bekommen hat bzw. die sich auf die Anzeige in der Zeitung hin gemeldet haben, den Fragebogen zusammen mit einem Rücksendeumschlag, einem Begleitschreiben und einer Instruktion zuschickt.

Insbesondere wenn man Experimente durchführt, ist es wichtig, den Versuchsablauf penibel zu dokumentieren. Eine gute Protokollierung hilft dabei, im Nachhinein unerwartete Schwierigkeiten oder Probleme bei der Versuchsplanung bzw. der Versuchsdurchführung zu erkennen oder diejenigen Versuchspersonen zu identifizieren, die aus bestimmten Gründen (z. B. mangelnde Motivation, Hypothese korrekt erraten, Probleme bei der Datenspeicherung etc.) von der Datenanalyse ausgeschlossen werden müssen.

Arbeitet man mit standardisiertem Versuchsmaterial (wie z. B. einem Fragebogen, einem Test oder einer computergesteuerten Apparatur), ist es sinnvoll, eine Zwischenanalyse der Daten vorzunehmen und zu überprüfen, ob es Schwierigkeiten gibt (z. B. zu schwere oder missverständliche Fragen), die es im Extremfall nötig machen, den Versuch abubrechen.

Auch bei der Versuchsdurchführung ist die Einhaltung ethischer Standards essenziell (vgl. die ethischen Richtlinien der DGPs und des BDP). Hierzu gehören v. a.:

- ▶ die Aufklärung der Versuchspersonen über potenzielle Risiken vor Beginn des Experiments,
- ▶ der Hinweis darauf, dass die Teilnahme an der Untersuchung freiwillig ist und jederzeit ohne Angabe von Gründen abgebrochen werden kann,
- ▶ die Zusicherung von Vertraulichkeit bei der Aufbereitung und Auswertung der Daten,

- ▶ das Einholen einer Einwilligungserklärung seitens der Versuchspersonen,
- ▶ eine ausführliche und lückenlose Aufklärung der Versuchspersonen über den Zweck der Untersuchung spätestens nach Abschluss der Untersuchung.

2.2.6 Auswertung der Daten

Sobald die Erhebung abgeschlossen ist, kann die Auswertung beginnen. Als Erstes müssen die erhobenen Informationen in einer Weise aufbereitet werden, die sich zur Auswertung eignet. In den meisten psychologischen Untersuchungen werden die erhobenen Informationen durch numerische Codierung in Zahlen übersetzt. In unserem Beispiel könnte die Bejahung der Frage nach normabweichenden Verhaltensweisen mit der Zahl 1, die Verneinung mit der Zahl 0 codiert werden. Antworten auf die Frage nach der Häufigkeit normabweichender Verhaltensweisen in einem definierten Zeitraum könnte man zahlenmäßig übernehmen. Den Zeitraum bis zur letzten Übertretung einer Norm könnte man in Abschnitte einteilen (z. B. 1 Monat, 3 Monate, 6 Monate, 1 Jahr, 3 Jahre usw.) und für jeden Abschnitt eine Zahl festlegen (1, 2, 3, 4, 5 usw.) oder aber den Zeitraum in der Anzahl der vergangenen Monate codieren (1, 3, 6, 12, 36 usw.). Die Art der Codierung hängt davon ab, welche Auswertungsmethode angewendet werden soll. Um Fehler bei der Übertragung zu vermeiden, empfiehlt es sich, einen Codierplan zu erstellen.

Datenmatrix: Darstellung der Daten

Die Speicherung der codierten Informationen erfolgt heute in elektronischen Datenbanken. Meistens werden die Daten in Matrixform geordnet und dargestellt. Diese Form ist übersichtlich und entspricht der Struktur, die viele Auswertungsmethoden voraussetzen. Eine Matrix ist die Anordnung von Zahlen in Tabellenform. Die Tabelle ist durch die Anzahl der Zeilen und Spalten definiert. Einer bewährten Konvention entsprechend schreibt man in der Psychologie die codierten Informationen über eine Person nebeneinander. Personen stellen in einer Datenmatrix also die Zeilen dar. Die Spalten entsprechen der Art der Information, die erhoben wurde, in unserem Beispiel also den einzelnen Fragen des Fragebogens. In einer Zelle der Matrix steht dann die numerisch codierte

Antwort, die eine Person auf eine bestimmte Frage des Fragebogens gegeben hat.

Deskriptivstatistik: Beschreibung der Daten

Die eigentliche Auswertung besteht darin, dass die Zahlen, die in der Datenmatrix stehen, nach bestimmten Regeln kombiniert und zusammengefasst werden. In der Regel beginnt man mit einer einfachen deskriptiven (beschreibenden) Analyse der Daten. In unserem Beispiel wird man sich vielleicht zunächst dafür interessieren, wie häufig bestimmte Verhaltensweisen vorkommen. Um dies zu ermitteln, könnte man die Summe der Spalten ermitteln, in denen die Antworten der Probanden auf die Frage stehen, wie oft sie eine bestimmte Norm in einem definierten Zeitraum verletzt haben, wie oft sie z. B. in den vergangenen fünf Jahren Steuern hinterzogen und Versicherungen betrogen haben. Die Summe der jeweiligen Spalte sagt aus, wie häufig das entsprechende Verhalten in der Gruppe insgesamt vorkam.

Zentrale Tendenz. Dividiert man diese Summe durch die Zahl der Probanden, ergibt sich die durchschnittliche Häufigkeit der jeweiligen Verhaltensweise. Dieser Durchschnittswert beschreibt die Gruppe insgesamt, nicht mehr eine einzelne Person. Man spricht von einer zentralen Tendenz.

Streuung. Weiterhin könnte man in einer ersten Analysephase auch die Frage klären, wie stark sich Personen in der Häufigkeit, mit der sie ein bestimmtes normabweichendes Verhalten in einem definierten Zeitraum zeigen, voneinander unterscheiden. Möglicherweise gibt es eine größere Gruppe, die das Verhalten nie zeigt, eine zweite größere Gruppe, die das Verhalten selten zeigt, und eine kleine Gruppe, die das Verhalten sehr häufig zeigt. Analysen dieser Art intendieren die Beschreibung der Verteilung oder Streuung von Verhaltensweisen in der untersuchten Gruppe.

Kovariation. In einem weiteren Analyseschritt könnte man sich der Frage zuwenden, ob zwei Verhaltensweisen (z. B. Steuerhinterziehung und Versicherungsbeitrag) im Sinne der Dispositionshypothese etwas miteinander zu tun haben. Um dies zu ermitteln, könnte man auszählen, wie viele Personen beide Verhaltensweisen (Fall 1), nur eine der beiden Verhaltensweisen (Fall 2) oder keine der beiden Verhaltensweisen