

Markus Pospeschill

Empirische Methoden in der Psychologie



Reinhardt

UTB

Markus Pospeschill

Empirische Methoden in der Psychologie

Mit 41 Abbildungen und 95 Übungsfragen

Ernst Reinhardt Verlag München Basel

PD Dr. *Markus Pospeschill* lehrt und forscht als Akademischer Direktor im Fach Psychologie an der Universität des Saarlandes Saarbrücken u. a. in den Bereichen Methodenlehre, Forschungsmethoden und Psychodiagnostik.

Vom Autor außerdem im Ernst Reinhardt Verlag erschienen:

„Testtheorie, Testkonstruktion, Testevaluation“ (UTB-M 978-3-8252-3431-7) und „Psychologische Diagnostik“ (zusammen mit Frank Spinath, UTB-basics 978-3-8252-3183-5).

Hinweis

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnungen nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

UTB-Band-Nr.: 4010
ISBN 978-3-8252-4010-3

© 2013 by Ernst Reinhardt, GmbH & Co KG, Verlag, München

Dieses Werk einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne schriftliche Zustimmung der Ernst Reinhardt, GmbH & Co KG, München, unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen in andere Sprachen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Printed in Germany
Einbandgestaltung: Atelier Reichert, Stuttgart
Covermotiv: © olly/fotolia.com
Satz: Da-TeX Gerd Blumenstein, Leipzig

Ernst Reinhardt Verlag, Kemnatenstr. 46, D-80639 München
Net: www.reinhardt-verlag.de E-Mail: info@reinhardt-verlag.de

Inhalt

Hinweise zur Benutzung dieses Lehrbuchs	9
Prolog	10
1 Methoden, Methodologie, Empirie	14
1.1 Methode	14
1.2 Methodologie	16
1.3 Empirie	17
1.4 Empirische Daten und Variablen	18
1.5 Empirische Fragestellung und Hypothese	19
1.6 Theorien, Gesetze, Paradigmen	23
1.7 Grenzen empirischer Forschung	29
2 Forschungsprozess	34
2.1 Untersuchungsidee und Thema	35
2.2 Wissenschaftliche und ethische Kriterien	37
2.3 Problempräzisierung	39
2.4 Definition von Begriffen	44
2.5 Indikatoren und Operationalisierung	50
2.6 Messen und Skalieren	56
2.7 Festlegung von Untersuchungsart und -objekten	60
2.8 Planung, Durchführung, Auswertung	64
2.9 Empirischer Bericht, Gutachten	66
3 Evaluation	74
3.1 Arten systematischer Erfolgskontrollen	74
3.2 Studienarten	80
3.3 Kosten-Nutzen-Analyse	86
3.4 Stichprobenauswahl	87
3.5 Durchführung und Auswertung	88
4 Methoden der Datenerhebung	91
4.1 Auszählen und Bilden von Indizes	91
4.2 Rangbildung und Paarvergleich	94
4.3 Ratingskalen	104
4.4 Testskalen	111
4.5 Schriftliche vs. mündliche Befragung	118

4.6	Formen wissenschaftlicher Beobachtung	124
4.7	Psychophysiologische Messung	126
4.8	Qualitative Interviews und Beobachtung	129
4.9	Multimethodale Methode	133
5	Hypothesengenerierende Untersuchungsformen	137
5.1	Voruntersuchung und Vortestung	137
5.2	Theoriebasierte Explorationsstudien	138
5.3	Methodenbasierte Explorationsstudien	140
5.4	Empiriebasierte Explorationsstudien	142
6	Populationsbeschreibende Untersuchungsformen	144
6.1	Zufallsstichprobe und Repräsentativität	144
6.2	Punkt- und Intervallschätzung	149
6.3	Probabilistische Stichprobentechniken	150
6.4	Nicht-probabilistische Stichproben	153
7	Hypothesenprüfende Untersuchungsformen	156
7.1	Signifikanztests und damit verbundene Probleme	156
7.2	Zusammenhangs-, Unterschieds-, Veränderungshypothesen	160
7.3	Prognostische Hypothesen und die Messung von Differenzen	177
7.4	Poweranalyse	180
7.5	Effektgrößenbestimmung	181
7.6	Schätzung optimaler Stichprobenumfänge	184
7.7	Strategien zur Vereinheitlichung von Effektgrößen	186
8	Einzelfallprüfende Untersuchungsformen	191
8.1	Problematik von Einzelfallstudien	191
8.2	Statistische Verfahren für Einzelfalldesigns	195
8.3	Statistische vs. visuelle Analyse	201
8.4	Randomisierungs- und Permutationstests	210
8.5	Verallgemeinerte Prinzipien des Randomisierungstests	225
8.6	Einzelfalldiagnostik	229

9	Besondere Probleme und Herausforderungen	249
9.1	Parametrische vs. verteilungsfreie Tests	249
9.2	Zusammenfassung statistischer Einzelentscheidungen	250
9.3	α -Fehler-Adjustierung	254
9.4	Bootstrap-Methode	258
9.5	Exakte Tests	259
 Literatur		265
Sachregister		269

Online-Materialien zum Buch stehen auf der Homepage des Ernst Reinhardt Verlags unter <http://www.reinhardt-verlag.de> sowie auf der UTB-Homepage unter <http://www.utb-shop.de/> zum Download zur Verfügung.

Hinweise zur Benutzung dieses Lehrbuchs

Zur schnelleren Orientierung werden in den Randspalten Piktogramme benutzt, die folgende Bedeutung haben:



Begriffserklärung oder Definition



Kritik oder eine wichtige Anmerkung



(Fall-)Beispiel



Übungsfragen am Ende der Kapitel

Prolog

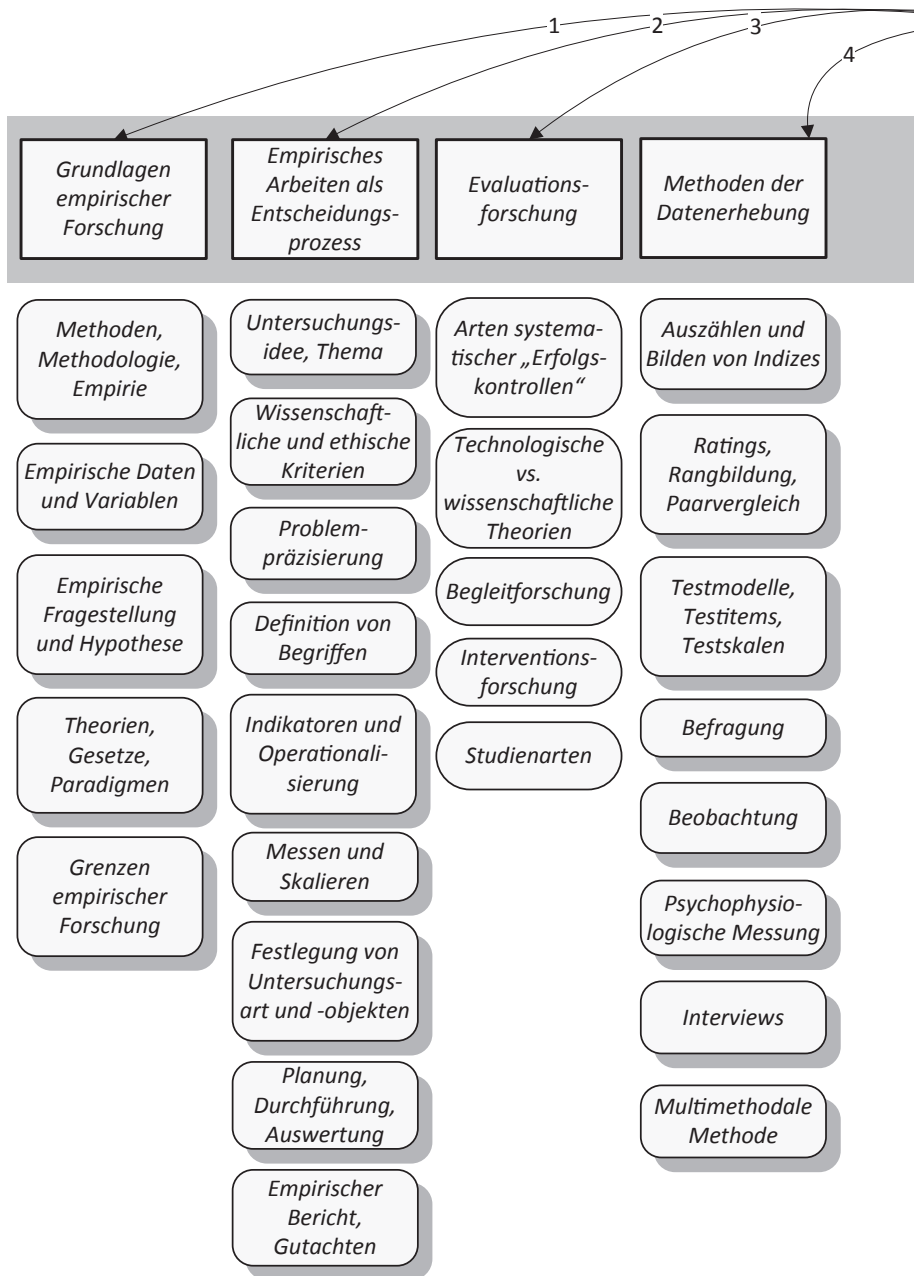
Die Auffassungen darüber, wie durch Auswertung systematischer Erfahrungen wissenschaftlich und praktisch bedeutsame Erkenntnisse abgeleitet werden, sind durchaus divergent. Dies gilt zuweilen selbst für den Kontext empirischen Arbeitens. Umso wichtiger ist es, die eigene Forschungsausrichtung und Arbeitsstrategie mit der nötigen Expertise, Sorgfalt und Gewissenhaftigkeit zu betreiben, ohne anderen Auffassungen (von Wissenschaft oder praktischer Arbeit) dabei ihre Existenzberechtigung abzusprechen.

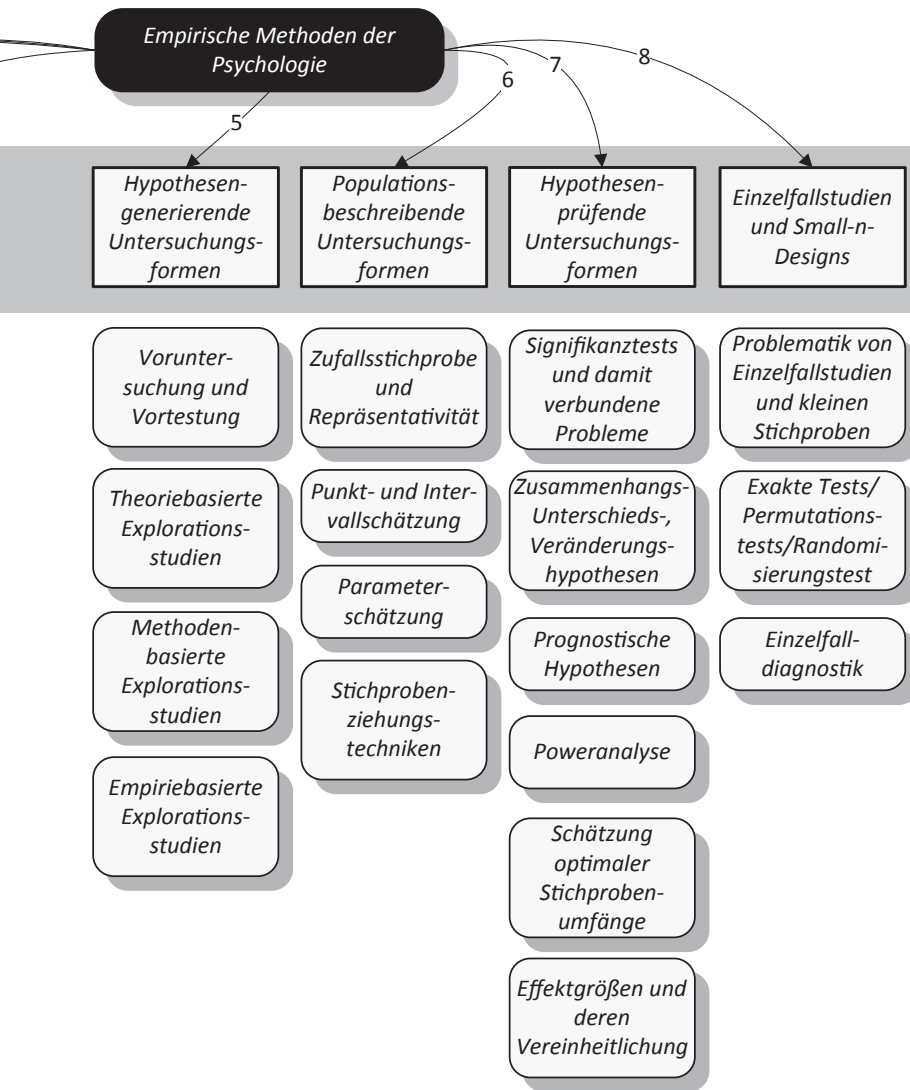
Darüber hinaus gilt, dass professionelles Handeln insbesondere im Human- und Sozialbereich vor allem methodisch geleitet sein muss. Damit kommen der Planung und Vorbereitung von entsprechenden Projekten besondere Bedeutung zu, um beispielsweise auch gegenüber anderen die einzelnen Schritte eines Planungsprozesses, der konkreten Vorgehensweise und der resultierenden Ergebnisse offenlegen und begründen zu können.

Die hier vertretene Auffassung orientiert sich wesentlich am Kritischen Rationalismus, einer wissenschaftstheoretischen Position von Karl R. Popper (1994) die insbesondere in den empirischen Human- und Sozialwissenschaften besondere Berücksichtigung findet. Wie bei anderen wissenschaftstheoretischen Positionen gilt es, die damit verbundene Sicht von Wissenschaft (und realweltlichen Problemen) in ihren Stärken und Schwächen beschreiben, verstehen und bewerten zu können. Empirische Forschung ist dabei als Erfahrungswissenschaft darauf ausgerichtet, im Fokus einer Forschungsfrage relevante Informationen zu generieren, zu strukturieren, systematisch auszuwerten, zueinander in Beziehung zu setzen und schließlich unter Anwendung von Theorien zu interpretieren, zu bewerten und in den Forschungskontext einzuordnen. Für die professionelle Tätigkeit im Human- und Sozialbereich gilt dieses Vorgehen gleichermaßen.

Um diesem Verständnis näher zu kommen, wird diese Abhandlung daher zunächst das begriffliche Rüstzeug der empirischen Forschung erarbeiten (**Kap. 1**) und auf die Planungsschritte eines Untersuchungsprozesses übertragen (**Kap. 2**). Als dann werden verschiedene Settings von Datenerhebungen (**Kap. 4**) und speziellen Anwen-

dungsvarianten empirischer Forschungsprinzipien eingehend thematisiert (**Kap. 3, 5–7**). Den Abschluss bilden Techniken, die eine Bewertung bzw. Bedeutungszuweisung empirischer Befunde ermöglichen sollen (**Kap. 8**).





Gesondert behandelte Themen:



1 Methoden, Methodologie, Empirie

Einleitend werden im Folgenden zentrale Begriffe zum empirischen Arbeiten geklärt, u. a. (wissenschaftliche) Methode, (allgemeine und spezielle) Methodologie, Empirie und empirische Variable, Hypothese, Theorie sowie Gesetzesaussage.

1.1 Methode



Im allgemeinen Sinne ist eine **Methode** (abgeleitet von dem griechischen Wort „méthodos“, dass so viel wie ‚der Weg auf ein Ziel hin‘) ein mehr oder weniger genau beschreibbarer Weg, d.h. eine endliche Folge von mehr oder weniger konkreten Handlungsanweisungen oder strategischen Maximen, zur Realisierung eines bestimmten Zieles bzw. zur Lösung einer bestimmten Aufgabe

wissenschaftliche Methode

Eine *wissenschaftliche Methode* soll darüber hinaus an gewisse Regeln oder Systeme von Regeln gebunden sein, welche die einzelnen Schritte der Vorgehensweise festlegen. Diese sollten *kommunizierbar*, *lehrbar* und *intersubjektiv nachprüfbar* sein. Weiterhin sollten sie *normativ* und *präskriptiv* (vorschreibend) sein, d.h. ein Verstoß gegen die Regeln einer Methode kann oder soll Sanktionen zur Folge haben. Eine wissenschaftliche Methode liefert damit ein begründbares, kommunizierbares und nachprüfbares Wissen, das bestimmten wissenschaftlichen Kriterien (z. B. der Allgemeingültigkeit, der Systematisierbarkeit etc.) genügt (Sprung & Sprung, 1987).

Beispielsweise gibt es Methoden zur Planung und Durchführung von Experimenten, die u.a. die Bildung einer (oder mehrerer) Kontrollgruppe(n) vorschreiben, welche aus der gleichen Grundgesamtheit (oder Population) wie die Experimentalgruppe stammen muss (müssen). Diese Vorschrift ist kommunizier- und lehrbar und ihre Befolgung außerdem intersubjektiv kontrollierbar. Verstößt nun ein Experimentator gegen diese, indem er z. B. die Bildung einer Kontrollgruppe unterlässt oder diese aus einer anderen Population bildet als die Experimentalgruppe, dann liegt eine fehlerhafte Anwendung derjenigen Methoden vor, die zur Planung und Durchführung von Experimenten entwickelt

wurden. Als mögliche Sanktion kann man sich hier z. B. die Ablehnung eines Manuskriptes durch die Gutachter einer Fachzeitschrift oder die Zurückweisung einer Studienabschlussarbeit oder Dissertation vorstellen.

Unter Methoden, insbesondere wissenschaftlichen Methoden, sollen hier also solche Methoden verstanden werden, die als Regeln vorliegen, die kommunizier- und lehrbar, normativ und präskriptiv sowie intersubjektiv nachprüfbar sind. Wissenschaften (ebenso wie Professionen) zeichnen sich generell durch methodische Vorgehensweisen aus, oder sollten sich wenigstens dadurch auszeichnen (um auch als echte Profession mit wissenschaftlich fundiertem Wissen sowie einer gerechtfertigten Fach- und Sachautorität zu gelten).

Regeln

Erkenntnis- vs. Interventionsmethoden

Methoden können weiter differenziert werden nach *Erkenntnismethoden* (die im engeren Sinne der Beschreibung, Erklärung und Vorhersage von Verhalten und Erleben dienen sollen) und *Interventionsmethoden* (die der Prävention, Rehabilitation, Stabilisierung oder Veränderung von Verhalten und Erleben dienen sollen). Ferner können Methoden zur *Datengewinnung* (verschiedener Formen der Erhebung, Registrierung und Kodierung von Verhaltensdaten) und Methoden zur *Datenanalyse* (der systematischen Auswertung von Verhaltensdaten) unterschieden werden. Im Folgenden wird es primär um Erkenntnismethoden und Methoden zur Datengewinnung gehen (Sarris, 1990).

Methodik vs. Modell

Zur Klärung der Terminologie seien einige Abgrenzungen zum Methodenbegriff angemerkt:

begriffliche Abgrenzung

Unter **Methodik** kann man allgemein die Lehre von den Methoden verstehen.



Eine Methodik zu haben oder zu verfolgen kann aber bedeuten, dass jemand in einer Forschungsarbeit methodisch vorgeht, d. h. eine (oder mehrere) bestimmte Methoden anwendet. Die Psychologie unterscheidet dazu eine Beobachtungs-, eine Befragungs-, eine Test- sowie eine Experimentalmethodik.

Im Gegensatz geht man bei der Verwendung von *Modellen* davon aus, dass hier Abbilder ausgewählter Aspekte der Realität in vereinfachter Form zum Einsatz kommen. Allerdings wird der Begriff Modell in den Wissenschaften mit verschiedenen Bedeutungen verwendet. Man spricht von einem *Denkmodell*, wenn man eine noch nicht ganz präzise und wenig überprüfte theoretische Vorstellung über einen Gegenstand meint.

Modellrelation

Zwischen dem Modell und seinem Urbild besteht immer eine bestimmte Beziehung, die *Modellrelation*. Man kann von bestimmten Merkmalen des Modells auf bestimmte Merkmale des Urbildes schließen und umgekehrt. Es besteht somit eine *Isomorphierelation*, eine umkehrbar eindeutige Abbildung, bei der alle Relationen erhalten bleiben.

1.2 Methodologie

Unter *Methodologie* kann das Nachdenken über Methoden verstanden werden: Eine *allgemeine (fachübergreifende) Methodologie* betrifft dabei alle Wissenschaften, während eine *spezielle (fachspezifische) Methodologie* nur in Wissenschaften einer bestimmten Gruppe angewendet wird (z. B. den empirischen Wissenschaften und dort ggf. in den Naturwissenschaften). So gehören beispielsweise statistische und experimentelle Methoden zum Untersuchungsgegenstand einer speziellen Methodologie (dort z. B. die hypothetico-deduktive und die signifikanzstatistische Methodologie). Weiter wird nach den Aspekten, unter denen die Methodologie wissenschaftliche Methoden untersucht, unterschieden: Eine *deskriptive Methodologie*, die Methoden beschreibt, klassifiziert und die Präferenz für bestimmte Methoden feststellt, und eine *wertende Methodologie*, die Methoden hinsichtlich ihrer Leistungsfähigkeit, ihren Voraussetzungen, ihrer praktischen Durchführbarkeit und der Kontrollierbarkeit bewertet.

**Wissenschafts-
theorie**

Insbesondere die fachübergreifenden Methodologien sind auch Gegenstandsbereich der *Wissenschaftstheorie*, einer *Metawissenschaft*, um grundlegende Prinzipien, Ziele und Vorgehensweisen in einzelnen Wissenschaften festzulegen. Neben den allgemeinen Verfahrensweisen zur Gewinnung wissenschaftlicher Erkenntnisse spielen in der Wissenschaftstheorie aber auch die Möglichkeiten und Grenzen menschlicher Erkenntnis (als Gegenstand von *Erkenntnistheorien* oder *Epistemologien*) und Fragen des Aufbaus und der Struktur von Theorien, Begriffen und Annahmen (*Metatheorien*) eine wesentliche Rolle.

Bei den Erkenntnistheorien können vier Grundpositionen abgeleitet werden (Westermann, 2000): **Erkenntnistheorien**

- *Realismus*. Der Realismus nimmt eine vom Betrachter unabhängige Außenwelt an, die durch Wahrnehmung und Denken erkannt und in ihren Wesensbeziehungen verstanden werden kann. Nach dem Kritischen Realismus kann man sich durch Wahrnehmung der Außenwelt allerdings nur annähern, ohne sie vollständig zu erfassen.
- *Idealismus*. Der Idealismus konstatiert eine rein geistige Wirklichkeit und Aussagen über die Wirklichkeit als reine Vorstellung. Radikale Idealisten verneinen dabei die Existenz einer vom Betrachter unabhängigen Welt. Kritische Idealisten hingegen bestreiten, dass ein Betrachter diese Außenwelt erkennen kann.
- *Empirismus*. Bei Empiristen gilt die Sinneserfahrung als einzige bzw. wichtigste Quelle von Wissen und Erkenntnis. Mit dem Verstand lassen sich diese Erkenntnisse strukturieren und generalisieren.
- *Rationalismus*. Rationalisten gehen davon aus, dass sich die Wirklichkeit durch Denken erkennen lässt. Wissen muss sich daher auf den Verstand und die Vernunft gründen können.

1.3 Empirie

Unter **empirisch** (abgeleitet von dem griechischen Wort „empeiria“) wird allgemein ‚sich auf Erfahrung gründend‘ bzw. ‚der Erfahrung angehörend‘ verstanden. Empirie ist dabei nicht mit dem oben beschriebenen Empirismus zu verwechseln. Beim Empirismus handelt es sich um einen philosophisch-erkenntnistheoretischen Standpunkt, nach dem die einzige Quelle von Wissen und Erkenntnis die Erfahrung ist. Bei der Empirie hingegen handelt es sich um ein methodisches Vorgehen, Theorien und daraus abgeleitete Hypothesen an Erfahrungsdaten zu überprüfen.



Mit *Erfahrung* ist hier die Informationsgewinnung mit Hilfe erweiternder Apparaturen (z. B. Mess- oder Testinstrumente) zu verstehen. Empirische Forschungsmethoden sind demnach solche Methoden, die zur Datenbeschaffung und systematischen Auswertung (Informationsgewinnung) über Ausschnitte der Realität eingesetzt werden. Sie konkretisieren sich in bestimmten Datenerhebungsinstrumenten wie Ex-

Erfahrung

periment, Test, physiologische Messung, Beobachtung, Befragung, Interview, Inhaltsanalyse etc. Daneben spielen die Datenauswertungsinstrumente eine entscheidende Rolle, die sich der Methode der Statistik bedienen.

1.4 Empirische Daten und Variablen

Die Objekte empirischer Untersuchungen (z. B. Einzelpersonen, Personengruppen, Organisationen) werden hinsichtlich ihrer *Merkmalsunterschiede* untersucht. Diese werden mittels *Variablen* erfasst. Eine Variable dient dabei als ein Bezeichner (Symbol) für eine Menge von Ausprägungen eines Merkmals. Dabei kommt im Rahmen der empirischen Forschung der Erhebung, Analyse und Interpretation registrierter Merkmalsunterschiede eine besondere Bedeutung zu.

Funktion von Variablen

Merkmalsausprägungen entstehen, wenn qualitativen Merkmalen Zahlen zugeordnet werden. Aus ihrer Menge entstehen *Daten*. Die Strategie, um von Merkmalen zu Daten zu gelangen, wird als *Operationalisierung* bezeichnet. Dabei gibt es grundsätzlich verschiedene Formen einer angemessenen Operationalisierung für ein Merkmal. Operationalisierungen können sich zudem auch auf nicht direkt beobachtbare sogenannte *latente Merkmale* (Konstrukte) beziehen. Allgemein können Variablen sehr unterschiedliche funktionale Bedeutungen besitzen, u. a. als:

- *Abhängige oder unabhängige Variable* (AV, UV). Abhängige Variablen sind Mess-, oder Kriteriumsvariablen, während unabhängige Variablen als Einfluss-, Prädiktor-, Faktor-, oder Gruppierungsvariable bezeichnet werden.
- *Moderatorvariable*. Diese verändert den Einfluss einer UV auf eine AV.
- *Mediatorvariable*. Liegt vor, wenn eine UV über eine Drittvariable auf die AV wirkt.
- *Kontrollvariable* bzw. *Kovariate*. Dies sind vorsorglich erhobene Moderatorvariablen.
- *Störvariable*. Entsteht, wenn eine nicht erhobene Kontroll- bzw. Moderatorvariable angenommen wird, die Einfluss auf die Messung ausübt.

- *Stetige oder diskrete Variablen.* Variablen mit unendlichen vs. endlichen Ausprägungen.
- *Dichotome oder polytome Variablen.* Variablen mit zwei vs. mehrfach gestuften Ausprägungen.
- *Manifeste oder latente Variablen.* Beobachtbare vs. nicht-beobachtbare Variablen.
- *Exogene oder endogene Variablen.* Variablen ohne vs. mit direktem Einfluss durch andere Variablen (vergleichbar mit UV und AV).
- *Indikatorvariable.* Zumeist spezifische manifeste Variablen, die distinkte Informationen an latente Variablen liefern.
- *Residualvariable.* Eine Restvariable, zumeist nicht erklärter Varianzanteile.

1.5 Empirische Fragestellung und Hypothese

Eine Hypothese (aus dem Griechischen „hypóthesis“ für ‚Unterstellung‘ oder ‚Vermutung‘) ist eine Aussage, die, ohne mit Sicherheit als wahr erkannt zu sein, für bestimmte Zwecke (z. B. für wissenschaftliche Erklärungen oder Voraussagen) angenommen wird. Anders ausgedrückt: Eine Hypothese ist eine Annahme über einen realen Sachverhalt. Drei Kriterien sind dabei entscheidend (Stier, 1999):



- *Der empirische Gehalt.* Wissenschaftliche Hypothesen müssen sich demnach empirisch untersuchen lassen.
- *Die Generalisierbarkeit.* Wissenschaftliche Hypothesen müssen über ein singuläres Ereignis hinausgehend formuliert sein.
- *Die Falsifizierbarkeit.* Wissenschaftliche Hypothesen werden über *Konditionalsätze* („Wenn ... dann“ oder „Je ...desto“) formuliert und müssen potenziell widerlegbar sein.

Der Erfüllung dieser drei Kriterien sind allerdings Grenzen gesetzt. Der empirische Gehalt einer Hypothese kann eingeschränkt sein, wenn sich Phänomene der realen Welt nur indirekt z. B. über schlussfolgernde Indikatoren beobachten lassen. Gleiches gilt für die Generalisierbarkeit, wenn der Einsatz einer Interventionsmethode nur an Einzelfällen überprüft werden kann. Schließlich ist über Wahrscheinlichkeitsaussagen letztlich keine Falsifikation möglich, da hier lediglich stochastische Beziehungen zwischen Variablen angenommen werden, die Gegenbeispiele zulassen.



- Damit entfallen Aussagen, die grundsätzlich nicht falsifizierbar sind (gilt i. d. R. für „Es gibt ...“-Sätze) oder die immer zutreffen bzw. tautologisch sind (z. B. Kann-Sätze). Einer empirischen Überprüfung entziehen sich ebenso Annahmen über Ereignisse oder Merkmale, die weder direkt beobachtbar noch indirekt mit dahinterstehenden Merkmalen in Verbindung gebracht werden können (z. B. metaphysische Aussagen).

**Hypothese als
Konditionalsatz**

Im Kontext wissenschaftlicher Hypothesen stellen der Bedingungs- (Antezedenz) und Folge-Teil (Konsequenz) eines Konditionalsatzes Variablen (mit mehreren Ausprägungen) dar. Der Bedingungsteil wird dabei als unabhängige Variable (UV) und der Folge-Teil als abhängige Variable (AV) bezeichnet. Dabei ist vorauszusetzen, dass die UV des Bedingungsteils mindestens zwei Ausprägungen besitzt. Entsprechend der Ausprägungsstufe der UV sollen sich in der Folge die Variablen der AV ändern. Formal ausgedrückt bedeutet das, „Wenn x_1 , dann y_1 “, wobei x_1 und y_1 Ausprägungen der Variablen X und Y darstellen. Zwischen den Ausprägungen $y_1, y_2, \dots$ der Variablen Y werden zudem Beziehungen formuliert (z. B. $y_1 > y_2$ oder $y_1 \neq y_2$).

Je-desto-Sätze

Konditionalsätze, bei denen Bedingungs- und Folge-Teil quantitativ bzw. kontinuierliche Variablen enthalten, werden typischerweise als Je-desto-Sätze formuliert. Dadurch entstehen zahlreiche Wenn-dann-Sätze, deren Wenn-Teil (z. B. $x_1 > x_2 > x_3$) und Dann-Teil (z. B. $y_1 < y_2 < y_3$) in abgestufter Form zugeordnet sind.

Zusammengefasst bedeutet dies, dass mit einer wissenschaftlichen Hypothese eine mehr oder weniger präzise formulierte Beziehung zwischen zwei oder mehreren Variablen ausgedrückt wird, die generalisierbar (also für eine Population vergleichbarer Objekte oder Ereignisse geltend) sein soll.

**induktive vs.
deduktive Funktion**

In den Humanwissenschaften stellen Hypothesen allerdings wissenschaftliche Hilfsmittel (sogenannte Heuristiken) dar, um aus Theorien oder empirischen Gegebenheiten Behauptungen über kausale oder nicht kausale Beziehungen zwischen Variablen abzuleiten. Anfänglich recht allgemein gehalten werden sie im Zuge der weiteren Untersuchungsplanung präzisiert und zu einer Prädiktion des antizipierten Untersuchungsergebnisses verdichtet. Dabei können sie immer noch verschiedenen Zwecken dienen:

- Sie besitzen *induktive Funktion*, wenn sie zur Hypothesenerkundung eingesetzt werden.

- Sie besitzen *deduktive Funktion*, wenn sie zur Hypothesenprüfung eingesetzt werden.

Die Überführung der Theorie (bzw. theoretischen Hypothese) in die Empirie (empirische Hypothese) ist dabei mit der Notwendigkeit verknüpft, Entscheidungen zu treffen, wie die Theorie angewendet werden soll und welche empirischen Beobachtungen zur Überprüfung der Theorie geeignet sind. Über diese Konkretisierungen macht die Theorie im Allgemeinen keine Aussage.

Statistische Hypothesen

Statistische Hypothesen überführen die inhaltliche bzw. wissenschaftliche Hypothese schließlich in eine Form, mit der die intendierte Aussage quantifiziert wird. Dazu wird eine Annahme über die Verteilung einer oder mehrerer Zufallsvariablen (der Begriff bringt zum Ausdruck, dass die Untersuchungsobjekte zufällig ausgewählt wurden) oder eines oder mehrerer Populationsparameter dieser Verteilung postuliert. Die statistische Hypothese bezieht sich entsprechend nicht auf Stichprobenverhältnisse, sondern auf eine zugrunde liegende Grundgesamtheit bzw. auf spezifische Kennwerte dieser Grundgesamtheit.

Quantifizierung

Die Variablenbeziehungen sind dabei probabilistischer Art (also Wahrscheinlichkeitsaussagen), denen eine gewisse Schwankungsbreite (Varianz) zugestanden wird. Entsprechend sind wissenschaftliche Hypothesen, die auf Wahrscheinlichkeitsmodellen basieren, durch konträre Einzelfälle prinzipiell nicht falsifizierbar. Da die Hypothesen über Populationen nur anhand von Stichproben untersucht werden, sind sie auch nicht verifizierbar. Um dieses Problem zu „lösen“, wird eine Falsifizierbarkeit durch die Festlegung von Prüfkriterien (z. B. mittels des Konzeptes der statistischen Signifikanz) erzeugt. Die dabei zur Anwendung kommenden Signifikanztests entscheiden unter Maßgabe der Prüfkriterien über die Annahme (Nicht-Falsifikation) oder Ablehnung (Falsifikation) der statistischen Hypothese.

**probabilistische
Variablenbeziehung**

Kausaleffekte

Eine empirisch bestätigte Beziehung zwischen Variablen ist dabei nicht per se in Form einer Kausalbeziehung zu interpretieren. Kausale Interpretationen von Effekten ergeben sich nicht einfach durch eine Variablenbeziehung (zwischen UV und AV) in Form eines Konditio-

nalsatzes, sondern sind vielmehr vom Untersuchungsdesign (z. B. einem echten Experiment, **Kap. 8.1**) und inhaltlichen Erwägungen abhängig. Der Nachweis eines Kausaleffektes im Rahmen eines Experimentes impliziert, dass der Untersucher den vermuteten Effekt herstellen kann. Weiter ist typisch für eine Kausalhypothese, dass sich der Wenn-Teil (Bedingung, Ursache) und der Dann-Teil (Konsequenz, Wirkung) nicht sinnvoll austauschen lassen.

**mono- vs.
multikausaler Effekt**

Die Interpretation von Kausaleffekten wird in den Human- und Sozialwissenschaften erschwert, da die gemessenen Effekte in der AV nur selten durch die Wirksamkeit einer UV zustande kommen, sondern häufig durch die kombinierte Wirkung mehrerer unabhängiger Variablen. Im ersten Fall würde man von einem monokausalen (durch eine Ursache bedingt), im zweiten Fall von einem multikausalen Effekt (durch mehrere Ursachen bedingt) sprechen. Einfache Konditionalsätze schließen multikausale Erklärungen nicht aus, da eine im Wenn-Teil genannte UV nicht als einzige Wirkvariable behauptet wird.

Dennoch müsste man von einer monokausalen Hypothese fordern, dass die geprüfte UV in der Lage ist, die wahre Variabilität (Gesamtvarianz abzüglich der Fehlervarianz) der AV vollständig zu erklären. Multikausale Hypothesen setzen hingegen den Erklärungswert einzelner unabhängiger Variablen geringer an und versuchen durch Annahme mehrerer Einflussfaktoren (als UV) auch mehr Varianz bei den Merkmalsunterschieden (in der AV) aufzuklären. Die Varianz, die dabei unerklärt bleibt, wird auf Messfehler und Störvariablen (also Einflüsse, die weder durch die UV, noch durch Moderator- oder Kontrollvariablen erklärt werden können) zurückgeführt. Bei den statistischen Hypothesen kann zudem die Präzision der Formulierung schwanken:

- Am präzisesten sind sogenannte *Punkthypothesen*, z. B. wenn X um 1 % größer wird, dann steigt Y um 2 %.
- *Gerichtete Hypothesen* sind dagegen solche, die die Richtung einer vermuteten Beziehung angeben, also z. B. wenn X größer wird, dann wird auch Y größer.
- Dagegen vermuten *ungerichtete Hypothesen*, dass z. B. zwischen X und Y ein Zusammenhang besteht.

**Aussagen über
Signifikanztests**



Ergänzend sei an dieser Stelle auf ein häufiges Missverständnis hingewiesen: Entscheidungen über empirische Hypothesen beruhen auf den Ergebnissen von Signifikanztests. Dazu werden zwei Arten von Hypothesen formuliert. Mit der Nullhypothese wird eine Null-Differenz oder Differenz unterhalb eines Schwellenwertes postuliert (also kein Effekt). Mit der Alternativhypothese wird

eine Differenz größer Null oder oberhalb eines Schwellenwertes formuliert. Wenn es mit statistischen Mitteln nicht gelingt, die Nullhypothese abzulehnen (also einen Effekt nachzuweisen), wird diese aber dadurch nicht bestätigt. Die Zusammenfassung statistisch nicht-signifikanter Testergebnisse mit Phrasen wie „die Nullhypothese ist akzeptiert“ oder „die Nullhypothese wird beibehalten“ regen zu diesem Missverständnis an. Genau genommen kann über die nicht-abgelehnte (nonrejected) Nullhypothese nicht mehr gesagt werden, als dass sie nicht abgelehnt wurde bzw. dass der Test nicht zur Ablehnung (failed to reject) der Nullhypothese geführt hat. Auch können keine praktischen Folgerungen für die wissenschaftliche Fragestellung (die der Hypothese zugrunde liegt) gezogen werden, wenn die Nullhypothese nicht abgelehnt wurde.

Ein weiteres Problem stellt sich ein, wenn zahlreiche Variablen oder Skalen verwendet werden, ohne dass dazu konkrete Hypothesen vorliegen. Dieser Umstand verleitet zu einem Pseudo-Hypothesentesten, bei dem durch Inspektion der Daten ein Effekt entdeckt, daraufhin (nachträglich) eine Hypothese formuliert, ein Signifikanztest gerechnet und das Ergebnis dann als ‚Bestätigung‘ der Hypothese interpretiert wird. Man beantwortet dadurch gewissermaßen Fragen, die man nie zuvor gestellt hat. Da der Signifikanztest hier auf den bereits entdeckten Effekt zugeschnitten wurde, bleibt das Ergebnis wenig überraschend (es gleicht der Freude über die gefundenen Ostereier, die man selber zuvor versteckt hat).

Pseudo-Hypothesentesten

1.6 Theorien, Gesetze, Paradigmen

In der Wissenschaftstheorie (einer Methodologie) herrschen verschiedene Traditionen, die sich wesentlich in

Wissenschaftsposition

- eine *analytisch-nomologische Position* (*Kritischer Rationalismus* nach Popper) und
- eine *hermeneutisch-dialektische Position* (*Frankfurter Schule*)

unterteilen lässt. Die Entscheidung für eine bestimmte Wissenschaftsposition ist ein Werturteil. Basis für die empirische Forschung ist der Kritische Rationalismus. Der kritische Rationalismus folgt dabei dem Hauptprinzip: Alle Aussagen einer empirischen Wissenschaft müssen an der Erfahrung überprüfbar sein und sie müssen prinzipiell an der Erfahrung scheitern können. Aufgabe der Wissenschaft ist es, Erklärungen für empirische Phänomene zu liefern. Die hermeneutisch-dialektische Position hingegen charakterisiert Aussagen als dialektisch, deren Gültigkeit diskursiv auszuhandeln ist. Die Aufgabe besteht darin, bestimmte Zusammenhänge in ihrem Wesen zu verstehen.

Begriff und Aussage Mit diesem Abgrenzungskriterium zu den nicht-empirischen Wissenschaften ergeben sich zunächst drei Konsequenzen für die praktische empirische Arbeit:

1. Alle *Begriffe*, die in einer Aussage vorkommen, die empirisch überprüft werden soll, müssen sich auf die erfahrbare Realität beziehen, anders ausgedrückt: Die Begriffe müssen einen *empirischen* Bezug haben.
2. *Aussagen*, die empirisch überprüft werden sollen, müssen sich auf Sachverhalte beziehen, die prinzipiell erfahrbar sind, anders ausgedrückt: Die Aussagen müssen *insgesamt einen empirischen Bezug* aufweisen.
3. *Aussagen*, die empirisch überprüft werden sollen, müssen so formuliert werden, dass sie *prinzipiell widerlegbar* sind.

Aus diesen Kriterien resultiert u. a.:

1. Begriffe sind ausgeschlossen, wie sie z. B. in Märchen oder Mythen vorkommen, die keinen empirischen Bezug aufweisen.
2. Bestimmte Aussagen werden ausgeschlossen, die sich unserem Zugriff entziehen, z. B. Aussagen über ein früheres (vorgeburtliches) Leben. Prinzipiell erfahrbar meint aber, Begriffe zuzulassen, die eventuell momentan aus technischen Gründen unserer Erfahrung noch nicht zugänglich sind.
3. Alle Aussagen, die an der Erfahrung gar nicht scheitern können, sind ausgeschlossen, z. B. *analytisch wahre* Aussagen, wie die Sätze der Logik und Mathematik, die nichts über die Realität aussagen, d. h. keinen empirischen Bezug haben, *Tautologien* („Wenn der Hahn kräht auf dem Mist, ändert sich das Wetter, oder es bleibt, wie es ist.“) oder *Existenzsätze*.

B Genügt die folgende Aussage den genannten Kriterien?

Bei älteren Arbeitnehmern ist die Arbeitszufriedenheit größer als bei jüngeren Arbeitnehmern.

Offensichtlich beziehen sich alle verwendeten Begriffe, soweit sie außerlogische sind (als logische Begriffe bezeichnet man z. B. und, oder, nicht, wenn, dann usw.), auf die erfahrbare Realität (Arbeitnehmer, Arbeitszufriedenheit) und außerdem bezieht sich die ganze Aussage auf einen real erfahrbaren Sachverhalt (Arbeitszufriedenheit). Es liegt weder eine analytisch wahre Aussage, noch eine Tautologie, noch eine Existenzaussage vor und die Aussage ist

widerlegbar: Findet man z. B. einen jüngeren Arbeitnehmer, dessen Arbeitszufriedenheit mindestens gleich groß ist wie bei einem älteren Kollegen, dann ist die Aussage widerlegt oder falsifiziert.

Allaussagen vs. Aussagen mittlerer Reichweite

Derartige Aussagen (oder Sätze), die in ihrem räumlichen und zeitlichen Geltungsbereich nicht eingeschränkt sind, werden als *Allaussagen* (oder *Allsätze*) oder auch als *nomologische Aussagen* oder *Gesetzesaussagen* bezeichnet. Aus einsehbaren Gründen stellen sie ein wissenschaftliches Ideal dar. In den Human- und Sozialwissenschaften sind solche Gesetzesaussagen nicht bekannt. Hier beschäftigt man sich vielmehr mit *räumlich* und *zeitlich* eingeschränkten Aussagen (sogenannten *Aussagen mittlerer Reichweite*), z. B. „Im Saarland war die Arbeitszufriedenheit im Jahr 2013 bei älteren Arbeitnehmern größer als bei jüngeren Arbeitnehmern“ oder „In der Firma ... war die Arbeitszufriedenheit im letzten Jahr ...“.

**uneingeschränkter
Geltungsbereich**

Ein Allsatz, so scheint es, ist falsifiziert, wenn wenigstens ein Fall gefunden werden kann, welcher im Widerspruch zu ihm steht. Wie steht es nun aber mit der Verifikation von Allsätzen? Am Beispiel sieht man leicht, dass Allsätze nicht verifiziert werden können. Dazu müsste man sämtliche Arbeitnehmer auf dem ganzen Globus zu allen Zeiten in eine empirische Untersuchung mit einbeziehen, was sowohl praktisch als auch prinzipiell unmöglich ist.

Basissatzproblem

Allerdings stellt sich für die Falsifikation bzw. Verifikation von empirischen Aussagen das *Basissatzproblem* (Popper, 1994): Wie können Hypothesen oder Theorien anhand empirischer Beobachtungen überprüft werden, wenn diese ebenfalls nur hypothetisch sind? Dieses Problem ist darauf zurückzuführen, dass Aussagen niemals direkt mit der Realität konfrontiert werden können, sondern immer nur indirekt, d. h. mit Aussagen über die Realität.

Zurück zum Beispiel mit der Arbeitszufriedenheit: Um z. B. eine Befragung praktisch durchzuführen, sind viele Entscheidungen notwendig, u. a. die Konstruktion eines Fragebogens, die Form der Befragung festlegen (persönlich, postalisch, per Telefon), den Fragebogen gestalten, eine Stichprobe nach bestimmten Kriterien definieren, die Untersuchung durchführen. Als Ergebnis dieser Bemühungen resultieren Daten, die nun in der Auswertung mit der Aus-

B

sage zur Arbeitszufriedenheit konfrontiert werden. Die Aussage wird aber nicht direkt mit der Realität konfrontiert. Diese Konfrontation hat schließlich zwei alternative Ausgänge: Die Daten stehen entweder im Widerspruch zur Aussage oder sie sind damit vereinbar.

Falsifikation und Verifikation

Kann nun behauptet werden, die Aussage sei damit falsifiziert bzw. verifiziert? Offensichtlich wäre das nur dann möglich, wenn ‚bewiesen‘ werden könnte, dass die Daten ‚wahr‘ sind. Aber gerade das ist nicht möglich, denn dazu müsste ebenso bewiesen werden, dass die ‚richtigen‘ Fragen gestellt wurden, dass die Befragten ‚wahrheitsgemäß‘ geantwortet haben und dass die ‚richtigen‘ Personen befragt wurden. Der Nachweis einer Falsifikation oder Verifikation ist daher niemals möglich. Bei der Erhebung von Daten können sich vielfältige Fehler einstellen, die den Wahrheitsgehalt in unkontrollierter Weise einschränken können.

empirische Bewährung

Was kann daraus geschlussfolgert werden? Statt von der Verifikation oder Falsifikation einer Aussage zu sprechen, ist daher angebracht, von einer empirisch *bewährten* bzw. *bestätigten* oder einer empirisch *widerlegten* bzw. *nicht bestätigten* Aussage (Befund) zu sprechen. In beiden Fällen ist zudem daran zu denken, dass es sich immer nur um eine *vorläufige* Bestätigung oder Widerlegung einer Aussage handelt. Natürlich bedeutet es für eine bewährte Aussage nicht, dass sie endgültig wahr sein kann. Nur wissen wir das niemals sicher. Folgerichtig werden niemals empirische Beweise geführt, auch wenn dies gelegentlich zumeist aus Übersetzungsfehlern resultierend so scheint (z.B. beim Begriff *empirical evidence*). Als Mindestforderungen an die empirische Wissenschaft bleiben daher:

- intersubjektive Nachprüfbarkeit des Vorgehens;
- Offenlegung der Vorgehensweise beim Forschungsprozess;
- Darlegung und Begründung aller Entscheidungen;
- angemessene Anwendung methodischen Fachwissens;
- korrekte Anwendung anerkannter Auswertungsinstrumente.

Deterministische vs. stochastische Aussagen



Noch mal zurück zum Beispiel: Angenommen, eine empirische Untersuchung der Aussage „In der Firma ... war die Arbeitszufriedenheit im letzten Jahr bei älteren Arbeitnehmern größer als bei jüngeren Arbeitnehmern“ führe zu dem Resultat, dass diese auf alle Arbeitnehmer der

Firma zutraf mit einer einzigen Ausnahme. Soll man nun wegen einer einzigen Ausnahme die Aussage als widerlegt ansehen? Das wäre wohl kaum als vernünftig zu bezeichnen. Man wird sich vielmehr die Frage stellen müssen, ob die Aussage selbst nicht zu hart formuliert war, nämlich als deterministische Aussage.

Das hat zur Konsequenz: Immer, wenn ein Arbeitnehmer der Firma älter ist als ein Kollege, dann ist seine Arbeitszufriedenheit auch höher als beim jüngeren Kollegen. An eine derartige Zwangsläufigkeit dürfte aber kaum jemand im Zusammenhang mit obiger Aussage denken. Vielmehr wird von einem weicheren Zusammenhang auszugehen sein, der auch Ausnahmen zulässt, etwa der Art: „im Allgemeinen“ oder „im Durchschnitt“ ist bei älteren Arbeitnehmern der Firma die Arbeitszufriedenheit größer als bei jüngeren Arbeitnehmern, oder auch: Die Wahrscheinlichkeit, mit seiner Arbeit zufrieden zu sein, ist für einen älteren Arbeitnehmer größer als für einen jüngeren. Derartige weichere Aussagen werden als stochastisch bezeichnet.

**stochastische
Aussage**

Theorien und Gesetze

Die Grundlage für Aussagen bzw. Hypothesen dieser Art liefern zumeist Theorien. Sie entstehen, wenn die relative Bedeutung von unabhängigen Variablen für die AV und die Beziehungen der unabhängigen Variablen untereinander ein erklärendes Netzwerk konstituieren. Ihre Gültigkeit bemisst sich dabei an ihrer Bewährung an realen Beobachtungen (auf deren Grundlage notwendige Daten erhoben werden).

Eine Theorie konstituiert sich allgemein als ein System von Aussagen, die sich auf einen bestimmten Gegenstandsbereich beziehen und logisch widerspruchsfrei sind, wobei die in den Aussagen auftretenden Begriffe definiert sein müssen.



Theorien können dabei einen unterschiedlichen Formalisierungsgrad aufweisen, bis hin zu einer Gesetzmäßigkeit (ein Sachverhalt, der unter eindeutigen und vollständig definierten Bedingungen stets gültig ist). Idealerweise enthält das Aussagensystem einer Theorie Axiome, d. h. grundlegende Postulate, die nicht aus anderen Sätzen abgeleitet werden können, aus denen unter Zuhilfenahme von logischen Transformations- und Ableitungsregeln weitere Aussagen und Theoreme deduzierbar sind. In der Psychologie und anderen Sozialwissenschaft-

Axiome

ten sind derart ausgearbeitete und formalisierte Theorien jedoch eher die Ausnahme.

Nichtsdestotrotz sollte eine gute Theorie die Befunde nicht nur beschreiben und erklären, sondern auch in einem erweiterten Geltungsbereich wertvoll sein, indem sie prognostisch über zukünftige Ereignisse und Entwicklungen Annahmen erlaubt.

Theorie oder Empirie

Im wissenschaftstheoretischen Kontext drängt sich dabei auch die Frage auf, was Priorität haben soll bzw. was wichtiger ist, die Theorie oder die Empirie. Aus empiristischer Perspektive sollte die Wissenschaft *induktiv* von Beobachtungen ausgehend zu Theorien gelangen, während aus rationalistischer Perspektive *deduktiv* aus Theorien Vorhersagen abzuleiten sind, die dann an Untersuchungsergebnissen zu überprüfen sind. Ein Blick in die Wissenschaftsgeschichte vieler Disziplinen zeigt allerdings, dass es unzutreffend wäre, ein Entweder-Oder zu postulieren, vielmehr dürfte von einer wechselseitigen Befruchtung zwischen Theorie und Empirie auszugehen sein. Naiv wäre es anzunehmen, empirische Forschung wäre voraussetzungslos, also völlig ohne Theorie, möglich. Mindestens rudimentäre Theorien bzw. Hypothesen sind unumgänglich.

Paradigma

Zeitliche Epochen sind zudem durch akzeptierte Vorgehensweisen (sogenannter *modus operandi*) gekennzeichnet, die ein Paradigma (aus dem Griechischen „*parádeigma*“ mit „*parà*“ für ‚neben‘ und „*deiknym*“ für ‚zeigen‘ oder ‚begrifflich machen‘) etablieren. Mit einem *Paradigma* etabliert sich eine wissenschaftliche Lehrmeinung. Ändert sich dieses Verständnis von Wissenschaftlichkeit, folgt zumeist ein *Paradigmenwechsel*.

Für die Psychologie bedeutet das, dass z. B. der Behaviorismus in den letzten Jahrzehnten an Einfluss verloren, kognitivistische Ansätze hingegen an Einfluss gewonnen haben. Aktuelle Dominanz hat vor allem eine neurobiologisch ausgerichtete Psychologie, während psychoanalytische Ansätze inzwischen weniger behandelt werden (auch wenn es gelegentlich Versuche gab, beide Positionen zu vereinen).

Werturteilsstreit

Schließlich ist wissenschaftliches Handeln auch in einen *Werturteilsstreit* eingebettet. Dabei geht es um die Frage, ob Werturteile (persönliche Meinungen z. B. zu Moral oder Politik) einen Einfluss auf die wissenschaftliche Forschung haben dürfen oder sollen, insbesondere, ob sie relevant sind oder sein sollen für die Akzeptanz von Theorien. Das würde bedeuten, dass dafür nicht (oder nicht nur) die Übereinstim-