

Vorlesungsskriptum: ALLGEMEINE WISSENSCHAFTSTHEORIE

Gerhard Schurz

Geschrieben 1988, überarbeitet 1996

Hinweis: Das mit * bezeichneten Kapitel 3.4 wird in der Vorlesung nicht behandelt. Es ist Thema der VL: Spezielle Wissenschaftstheorie.

Inhaltsverzeichnis:

1. Allgemeine Einführung

- 1.1 Fragen und Ziele der Wissenschaftstheorie. Allgemeine versus spezielle Wissenschaftstheorie.
- 1.2 Methode und Voraussetzungen der Wissenschaftstheorie
- 1.3 Das minimale gemeinsame erkenntnistheoretische Modell der Wissenschaft: fünf allgemeine Annahmen.
- 1.4 Die minimale gemeinsame Methodologie der Wissenschaft: vier Grundbestandteile.
- 1.5 Drei Arten der Induktion
 - 1.5.1 *Die methodische Induktion*
 - 1.5.2 *Die logische Induktion*
 - 1.5.3 *Die pragmatische Induktion*
- 1.6 Untergliederung wissenschaftlicher Disziplinen und Abgrenzung von "Wissenschaft".
- 1.7 Zur Frage der Wertfreiheit von Wissenschaft
- 1.8 Historischer Exkurs: Zur historischen Entwicklung der Wissenschaftstheorie
- 1.9 Literaturhinweise

2. Die logischen Grundlagen

- 2.1 Einteilung von Begriffsarten
 - 2.1.1 Einteilung nach dem logischen Typ
 - 2.1.2 Einteilung von Begriffen nach dem Inhaltstyp
 - 2.1.3 Einteilung der Begriffe nach ihrer Metrik bzw. nach ihrem Skalentyp
- 2.2 Klassifikation von Satzarten
 - 2.2.1 Klassifikation von Satzarten nach dem Inhaltstyp
 - 2.2.2 Logik und logische Wahrheit
 - 2.2.3 Bedeutungskonventionen und definitiorische Wahrheit
 - 2.2.4 Deskriptive versus präskriptive Sätze
 - 2.2.5 Klassifikation von Satzarten nach ihrem Allgemeinheitsgrad

- 2.4.6 Deduktive Beziehungen zwischen den Satzarten geordnet nach Allgemeinsgrad
- 2.4.7 Klassifikation von Satzarten nach ihrem methodologischen Status
- 2.5 Verifikation, Falsifikation, Bestätigung (Bewährung) und Schwächung
- 2.6 Literaturhinweise

3. Empirische Gesetze und ihre Überprüfung

- 3.1. Die Relevanzbedingung - deterministischer und statistischer Fall
- 3.2. Die empirische Überprüfung von Gesetzen auf vermutliche Wahrheit (Bestätigung vs. Falsifikation) und auf Relevanz
 - 3.2.1. Der deterministische Fall: Die Methode der Übereinstimmung (Bestätigung und Falsifikation) und des Unterschieds
 - 3.2.2 Der statistische Fall: die Methode der Konfidenzintervalle und der Signifikanztests
 - 3.2.3 Mögliche Fehlerquellen
- 3.3 Das Problem der Kausalität: Korrelation und Kausalität
 - 3.3.1 Versteckte Variablen
 - 3.3.2 Kausalrichtung
 - *3.4 Erklärung, Begründung und Voraussage: Der deduktiv-deterministische und der induktiv-statistische Fall
- 3.5 Literaturhinweise

4. Wissenschaftliche Theorien

- 4.1 Beobachtungsbegriffe, theoretische Begriffe und Zuordnungsgesetze
- 4.2 Exkurs: Erkenntnistheoretische Hierarchie von Begriffen
- 4.3. Strukturelle Bestandteile und methodologische Merkmale wissenschaftlicher Theorien - am Beispiel der Newtonischen Physik
- 4.4 Holismus der Theorienüberprüfung - am Beispiel der Newtonischen Physik
- 4. 5 Zweites Beispiel: Piagets kognitive Entwicklungspsychologie
- 4.6 Adäquatheit von Zuordnungsgesetzen bzw. Indikatoren in den Human- und Sozialwissenschaften
- 4. 7 Theorienbewertung, Theorienvergleich, und Theorienfortschritt
- 4.8 Die Bedeutung der empirisch-theoretisch-Unterscheidung für das Programm einer objektiven empirischen Wissenschaft
- 4.9 Erkenntnistheoretische Interpretation von Theorien: Instrumentalismus und Realismus
- 4. 10 Letztes Beispiel: Die Adorno-Milgram-Theorie des autoritären Charakters
- 4.11 Literaturhinweise

1. Allgemeine Einführung

1.1. Fragen und Ziele der Wissenschaftstheorie. Allgemeine versus spezielle Wissenschaftstheorie.

Im Zentrum der Wissenschaftstheorie stehen Fragen wie die:

Wesensbestimmung von Wissenschaft: Was ist Wissenschaft? Worin besteht sie? Was ist der Gegenstand wissenschaftlicher Erkenntnis? Was ist das Ziel wissenschaftlicher Erkenntnis? Was ist ihr Inhalt? Ihre Struktur (Statik) und ihre Entwicklung (Dynamik)?

Wissenschaftliche Methode: Wie funktioniert wissenschaftliche Erkenntnis? Was ist ihre Methode? Nach welchen Regeln geht sie vor? In welchem Ausmaß wird sie von Regeln geleitet?

Ziele und Leistungen: Was sind die Ziele der Wissenschaft, und was nicht? Was leistet wissenschaftliche Erkenntnis?; was kann sie leisten, und was nicht?

Allgemeinheiten und Besonderheiten: Welche gemeinsamen Merkmale aller Wissenschaftsdisziplinen gibt es, und wo liegen Unterschiede?

Die allgemeine Wissenschaftstheorie fragt nach jenen Elementen, die allen Wissenschaftsdisziplinen gemeinsam sind - also sie stellt Fragen wie:

Wie ist eine wissenschaftliche Sprache aufgebaut? Was sind die Regeln für die Korrektheit bzw. Gültigkeit eines wissenschaftlichen Argumentes? (Logik i.w.S.)

Was ist ein wissenschaftliches 'Datum' bzw. eine wissenschaftliche Beobachtung? (Empirie)

Was ist ein wissenschaftliches Gesetz, eine Theorie? (Strukturanalyse)

Was ist eine wissenschaftliche Voraussage, was eine wissenschaftliche Erklärung? (Anwendung von Theorien auf Erfahrung;)

Wie werden Gesetze und Theorien überprüft, bestätigt (bewährt) bzw. falsifiziert? (empirische Überprüfung)

Was sind die Kriterien für rationalen Theorienfortschritt? (dynamische Analyse)

- u.a.m. Die allgemeine Wissenschaftstheorie behandelt aber auch die Frage nach den *charakteristischen Unterschieden* einzelner Wissenschaftsarten. Z.B. wurde und wird immer wieder die Frage diskutiert, ob es einen grundsätzlichen Unterschied gäbe zwischen den Natur- und den Geisteswissenschaften. Kennzeichen der allgemeinen Wissenschaftstheorie ist es jedenfalls, daß sie alle Disziplinen ins Auge faßt. Sie versucht, ihre Gemeinsamkeiten festzustellen, zugleich aber mögliche

charakteristische Unterschiede aufzudecken.- Die spezielle Wissenschaftstheorie, besser, die speziellen Wissenschaftstheorien, sind dagegen auf einzelne Disziplin-gattungen bezogene Wissenschaftstheorien, die sich mit den grundlagentheoretischen Fragen spezieller Disziplinen (wie Physik, Biologie, Humanwissenschaften, usw.) befassen.

Die allgemeinsten Fragen der Wissenschaftstheorie sind:

- was ist (wissenschaftliche) Wahrheit?, und
- worin besteht (wissenschaftliche) Objektivität?

Bei diesen Fragen geht *Wissenschaftstheorie in Erkenntnistheorie* über, die sich mit diesen allgemeinen Fragen hinsichtlich der Natur von Erkenntnis überhaupt beschäftigt.

Mit der Auflistung dieser Fragen haben wir natürlich zugleich die primären Ziele der Wissenschaftstheorie angeführt. Ihre Ziele liegen in der Suche nach wahren (rechtfertigbaren, haltbaren) Antworten auf diese Fragen. Man kann diese Ziele auch die internen Ziele der Wissenschaftstheorie nennen.

Darüber hinaus hat die Wissenschaftstheorie, wie jede Wissenschaft, auch externe Ziele, die den Wert der Wissenschaftstheorie für andere Zwecke, ihre Anwendung auf andere Gebiete und Probleme betreffen. Man kann hierbei wieder zwischen wissenschaftsinternen und wissenschaftsexternen Anwendungsmöglichkeiten unterscheiden. Wie zählen jeweils einige Beispiele auf:

Wissenschaftsinterne Anwendung:

- (1) Entscheidungshilfen für kontroverielle Fragen in den Einzelwissenschaften.
- (2) Wegbereiter und Rahmendisziplin für neue Wissenschaftsdisziplinen, wie Kognitive Wissenschaft und Artificial Intelligence, oder Theorie der natürlichen Sprache.
- (3) Herausarbeitung interdisziplinärer Gemeinsamkeiten und Zusammenhänge zwischen verschiedenen Disziplinen. Zu diesem Zweck entwirft die Wissenschaftstheorie verallgemeinerte Modelle in einer übergreifenden, in alle 'Fachsprachen' übersetzbaren und möglichst präzisen Sprache - nicht zuletzt deshalb ist auch das Verfahren der logischen Explikation mithilfe logischer Modellsprachen für die Wissenschaftstheorie von großer Bedeutung.
- (4) Insofern die Wissenschaftstheorie dazu dient, Wissenschaft besser zu verstehen und zu durchschauen hat sie einen kognitiven Bildungswert, der sich für die Fähigkeit zu wissenschaftlichem Arbeiten immer wieder als sehr förderlich erweist.

Wissenschaftsexterne Anwendung:

(5) Unser praktisches Leben wird von Technik und Expertenwissen stark dominiert, ständig beruft man sich (z.B. in den Medien) auf zumeist unvollständig deklariertes Expertenwissen - hier ist ein kritisches und fundiertes Wissenschaftsverständnis sehr wichtig, um Seriöses von Unseriösem auseinanderhalten zu können.

(6) Zugleich sind einige technischen Anwendungen der Wissenschaft heute zu einer großen Gefahr geworden, und daher ist unsere Zeit durch ein kritisches Verhältnis zu Wissenschaft und Technik gekennzeichnet. Auch hier ist ein wirklich fundiertes Wissenschaftsverständnis eine äußerst wichtige Voraussetzung zu fundierter Kritikfähigkeit, also um bloße Vorurteile von wirklichen Problemen unterscheiden zu können.

(7) Im Zusammenhang mit obigem Problem wird heute auch immer wieder das Verhältnis von Wissenschaft zum seelisch-gefühlsmäßigen Bereich des Menschen, und in diesem Zusammenhang zu Kunst, Religion und Ethik diskutiert. Auch zur sachlich-kritischen Diskussion dieses Problems ist der Abbau von Vorurteilen, der durch ein wissenschaftstheoretisches Grundwissen ermöglicht wird, sehr vorteilhaft, nicht zuletzt, weil er dazu dienen kann, gesellschaftliche Sprachbarrieren zu überwinden.

1.2 Methode und Voraussetzungen der Wissenschaftstheorie

Es lassen sich zwei prima facie gegensätzliche Auffassungen zur Aufgabenstellung und Methode der Wissenschaftstheorie unterscheiden:

1. *die deskriptive Auffassung:* Wissenschaftstheorie hat die Aufgabe, zu sagen, was Wissenschaft ist. Dazu muß sie die faktisch vorliegenden Wissenschaften in ihrer gegenwärtigen Struktur sowie in ihrer historischen Entwicklung beschreiben und, soweit möglich, erklären.

2. *die normative Auffassung:* Wissenschaftstheorie hat die Aufgabe, zu sagen, was Wissenschaft sein soll. Dazu muß sie angeben, worin wissenschaftliche Rationalität besteht, d.h. aufgrund welcher Kriterien sich eine wissenschaftliche Hypothese rational rechtfertigen läßt.

Einschub: Vorklärung der Begriffe "deskriptiv" und "normativ": Ein Satz heißt deskriptiv, wenn er etwas über den Zustand der Realität behauptet, also wenn er sagt, was "ist". Ein Satz heißt normativ, wenn er sagt, was "sein soll", also nicht etwas über den Zustand der realen Welt behauptet, sondern einer ideale erwünschte Welt. - Die Unterscheidung zwischen deskriptiven und normativen Sätzen setzt offenbar einen minimalen Realismus (s.u.) voraus - es gibt eine subjektunabhängig existierende "Welt da draußen" (was noch nicht heißt, daß auch unsere Erkenntnis dieser Welt subjektunabhängig bzw. objektiv ist !). Deskriptive Sätze besagen etwas über diese Welt, und können insofern wahr oder falsch sein. Und zwar ist ein deskriptiver Satz, gemäß der sogenannten und weithin akzeptierten Korrespondenztheorie der Wahrheit wahr, wenn der von ihm behauptete Sachverhalt in der Realität tatsächlich besteht; er ist falsch, wenn der von ihm behauptete Sachverhalt nicht besteht. (Diese Korrespondenztheorie wurde vom Logiker Alfred Tarski in moderner Form präzisiert, sie geht historisch auf Aristoteles zurück). Da normative Sätze nichts über die reale Welt besagen, sondern etwas über das ethisch Wünschenswerte behaupten, sagt man üblicherweise von normativen Sätzen auch nicht, daß sie wahr oder falsch sind, sondern nur daß sie 'richtig' oder 'unrichtig', oder 'gültig' oder 'ungültig' sind. (Ende des Einschubs).

Die normative Auffassung ist historisch älter und gegenwärtig noch dominierend. Ihre Vertreter sind im Grunde alle 'modernen Klassiker', sowohl Popper wie der Wiener Kreis (am deutlichsten hier Viktor Kraft). Die deskriptive Gegenposition wurde dagegen massiv erst in jüngerer Zeit vertreten, z.B. von Stegmüller (1973, Band IV, 1. Teilband, S. 310). Welche Position ist 'die richtige'?

Die Argumente der Normativisten gründen sich auf die bekannte, auf von Hans Reichenbach zurückgehende Unterscheidung zwischen Entdeckungszusammenhang und Rechtfertigungszusammenhang (context of discovery versus context of justification), bzw. wie wir verallgemeinert sagen wollen: die Unterscheidung zwischen *Entstehungszusammenhang* und *Begründungszusammenhang* (resp. Bewertungszusammenhang) wissenschaftlicher Erkenntnis. Der Entstehungszusammenhang betrifft die Frage, wie wissenschaftliche Erkenntnis faktisch entsteht; der Begründungszusammenhang die Frage, wie sie sich wissenschaftliche Erkenntnis rational ausweisen, als vermutlich wahr oder empirisch bewährt begründen läßt. - Die Normativisten argumentierten nun zumeist so: der wissenschaftliche Wert einer Aussage hängt nur von ihrem Begründungszusammenhang, nicht von ihrem Entstehungszusammenhang ab. Daher solle sich die Wissenschaftstheorie nur mit dem Begründungs-, nicht aber mit dem Entstehungszusammenhang beschäftigen (letzterer ist Sache der Wissenschaftspsychologie, -soziologie, -geschichte, nicht

aber ihrer -theorie).

Es ist nun besonders wichtig zu sehen, worin der Fehlschluß dieser Argumentation liegt. Es ist nämlich ganz *richtig*, daß der wissenschaftliche Wert einer Aussage allein vom ihrer rationalen Begründungszusammenhang abhängt (insofern Wissenschaft rational und objektiv zu sein beansprucht). Beispielsweise ist es ganz irrelevant für den wissenschaftlichen Wert einer Theorie, ob ihr Erfinder ihr im Traum gewahr wird (wie Kekulé den Benzolring erträumte), oder ob er sie durch systematische Variation bekannter Theorien gewinnt; und ebenso irrelevant ist es, ob ihr Erfinder damit irgendwelche politischen Ziele verband, und wenn ja, welche. All das könnte im Entstehungszusammenhang eine Rolle gespielt haben, ist jedoch für den Begründungszusammenhang irrelevant; ja muß irrelevant sein, solange Wissenschaft nur am Ziel der Wahrheit, an nichts sonst, ausgerichtet sein soll. (Wohlgemerkt machen wir damit nicht die deskriptive Behauptung, daß die Wissenschaftler in ihren Begründungen sich faktisch immer am Wahrheitsziel orientieren. Wir machen die normative Behauptung: weil Wissenschaft letztlich auf Wahrheit zielt, darf für den Wahrheitsanspruch ihrer Theorien nur der Begründungszusammenhang eine Rolle spielen. Näheres unten).

Dennoch folgt daraus keineswegs, daß die Wissenschaftstheorie sich nicht mit dem Entstehungszusammenhang zu beschäftigen braucht. Man muß nur folgende Frage stellen, um die Unhaltbarkeit dieses Schlusses zu erkennen: wie *gelangen* denn die Normativisten zu den von ihnen behaupteten wissenschaftstheoretischen Normen und Regeln? Wenn sie diese nicht schon aus der Analyse der faktischen Wissenschaften gewinnen, *woher denn?* Wie wir schon aus Kap. 1.2 wissen, lassen sich solche Regeln nicht durch Intuition gewinnen - wie die Geschichte zeigt, entstand die Wissenschaftstheorie immer nur in Auseinandersetzung mit den faktischen Wissenschaften. Die Regeln der Wissenschaft lassen sich ebenso nicht aus bloßer Logik schöpfen (wie die Entwicklung des logischen Empirismus deutlich zeigte). Wir kommen daher zu folgendem Resultat: Obwohl es die Aufgabe der Wissenschaftstheorie ist, den Begründungszusammenhang von Erkenntnis, also die Kriterien ihrer Rationalität zu analysieren, kann sie das doch nur tun aufgrund Analyse der faktischen Wissenschaft in ihrer historischen Entwicklung.

Obwohl sich schon Carnap, Hempel und Popper an faktischen Wissenschaften - vorallem an Standardlehrbüchern der Naturwissenschaften - orientierten, zeigten doch die wissenschaftshistorischen Arbeiten von Kuhn (1976), daß viele ihrer Vorstellungen viel zu simpel waren, um reale Wissenschaft zu erfassen. In gewisser Weise wurde die Wissenschaftstheorie durch den von Kuhn ausgelösten "Einbruch der Wissenschaftsgeschichte" aufgeschreckt. Man bemerkte, daß die Wissenschaftstheorie sich in noch viel intensiverem Ausmaß als bisher mit den faktischen Wissenschaften und ihrer Entstehung beschäftigen müßte. (Die Polemik

von Paul Feyerabend 1976, derzufolge sich die Wissenschaftler besser nicht an den von der Wissenschaftstheorie aufgestellten 'Normen' orientieren sollten, sondern der Devise des "Anything Goes" huldigen sollten - was, nüchtern betrachtet freilich ein Freispruch für Ideologien jeglicher Art ist, der das Problem überhaupt nicht vorwärtsbringt - kam so gesehen nicht von ungefähr). Als Reaktion darauf proklamierten Stegmüller und andere jüngeren Autoren, Wissenschaftstheorie solle sich überhaupt auf deskriptive Analyse der Wissenschaften beschränken. Auch diese Position scheint uns allerdings übertrieben. Die Frage nach den Kriterien wissenschaftlicher Rationalität muß im Zentrum der allgemeinen Wissenschaftstheorie bleiben - das ist ja die Frage, die die allgemeine Wissenschaftstheorie als Disziplin zusammenhält (!). Die bloße empirische Analyse von Erkenntnis wird z.B. auch von der Wissenschaftsgeschichte, -soziologie und -psychologie, oder auf der ontogenetischen (individualgeschichtlichen) Ebene z.B. von der Kognitionspsychologie geleistet. Wissenschaftstheorie ist aber nicht identisch mit einer dieser Disziplinen, auch nicht mit einem Konglomerat davon. Sie unterscheidet sich von diesen rein empirisch-deskriptiven Disziplinen dadurch, daß sie sich um das Problem der Definition und Rechtfertigung wissenschaftlicher Rationalität bemüht. Kurz gesagt: *die Wissenschaftstheorie ist somit eine Disziplin, die sowohl deskriptive wie normative Bestandteile enthält.*

Die Methode der Wissenschaftstheorie läßt sich am besten als *rationale Rekonstruktion* bezeichnen. Zunächst einmal geht die Wissenschaftstheorie von folgender obersten *erkenntnisinternen Norm* aus: das allgemeine Erkenntnisziel der Wissenschaft ist die *Findung wahrer gehaltvoller Sätze* (vgl. ebenso Weingartner 1978, S. 46). D.h., die Sätze müssen nicht nur wahr, sondern auch gehaltvoll sein, wobei 'gehaltvoll' zumeist "empirisch gehaltvoll" bedeutet, und darüber hinaus auf die spezifischen Probleme, die gelöst werden sollen, bezogen ist.

Das Ziel der "wahren gehaltvollen Aussage" besagt allein jedoch noch überhaupt nichts. Ganz unabhängig von seinem 'Denkstil' würde sich wohl jeder Mensch darauf berufen, nur daß dabei jeder unter "gehaltvoller Wahrheit" etwas anderes verstehen mag. Das Ziel gewinnt erst Inhalt durch die zweite, vom Wissenschaftstheoretiker vorausgesetzte Komponente, ein *minimales gemeinsames erkenntnistheoretisches Modell* der Wissenschaft, das Begriffen wie 'Wahrheit' etc. erst Sinn verleiht. Wir werden in Abschnitt 1.4.1. versuchen, dieses minimale und insofern auch allgemeinste erkenntnistheoretische Modell der Wissenschaften, so wie es von der modernen, mit den empirischen Wissenschaften mitgewachsenen Wissenschaftstheorie in großer Mehrheit vertreten wird, zu explizieren. Hier sei dieses minimale erkenntnistheoretische Modell nur in Stichpunkten umrissen: es besteht aus (1) einem minimalen Realismus (im bereits oben angesprochenen Sinn), (2) der Einstellung der Fallibilität (Fehlbarkeit) allen wissenschaftlichen Wissens, (3)

dem Streben nach Objektivität, (4) die Verwendung möglichst präziser logisch-mathematischer Begriffe und Methoden, und (5) einem minimalen Empirismus in Form einer noch zu erläuternden empirischen Überprüfungsforderung für nicht schon logisch-mathematisch wahre Behauptungen.

Es sei schon hier erwähnt, daß über dieses in fünf Punkten dargestellte allgemeine und minimale erkenntnistheoretische Modell in der gegenwärtigen Wissenschaftstheorie nicht überall Einigkeit besteht. Dennoch wird dieses Modell, in eventuell unterschiedlicher Ausformung oder Gewichtung der einzelnen Punkte, von der Wissenschaftstheorie normalerweise vorausgesetzt (und nur in Ausnahmeständen hinterfragt, s.u.). Die eigentliche Tätigkeit der Wissenschaftstheorie besteht in schon erwähnten Methode der rationale Rekonstruktion, die nun als folgende Aufgabenstellung präzisiert werden kann:

Entwickle anhand faktisch vorliegender wissenschaftlicher Erkenntnissysteme (der Geschichte oder der Gegenwart) verallgemeinerte und möglichst logisch präzisierte Modelle von wissenschaftlicher Erkenntnis, die einerseits auf die faktische wissenschaftliche Erkenntnis - zumindest in ihren Musterbeispielen (s.u.) - zutreffen, andererseits sich auch unter Voraussetzung des allgemeinen Erkenntnisziels und des allgemeinen und minimalen erkenntnistheoretischen Erkenntnismodells rechtfertigen lassen, und zwar dadurch, daß sich die (disziplinspezifischen) Realisierungen dieser Modelle als optimale Mittel erweisen, unter Voraussetzung der minimalen erkenntnistheoretischen Annahmen das allgemeine Erkenntnisziel zu erreichen.

Man kann die faktischen wissenschaftlichen Erkenntnissysteme das deskriptive Korrektiv und das vorausgesetzte allgemeine Erkenntnisziel plus das minimale Erkenntnismodell das normative Korrektiv der Wissenschaftstheorie nennen (wobei das "normativ" hier im weiten Sinn von "normativ-methodologisch" verwendet wird, denn nur das Erkenntnisziel selber ist im engen Sinn ein Normsatz, während die Annahmen des minimalen Erkenntnismodells deskriptive Sätze sind, die nur indirekt normativen Charakter haben, insofern sie selber allgemeine Mittelthesen bezüglich der Erreichung des Erkenntnisziels sind, die normalerweise vorausgesetzt werden). Die Tätigkeit der Wissenschaftstheorie besteht darin, aus der deskriptiven Basis verallgemeinerte Modelle zu abstrahieren, die zugleich dem normativen Korrektiv entsprechen. Das sagt ja auch schon der Begriff der rationalen Rekonstruktion: man zeichnet etwas nach und erweist es dadurch *zugleich* als rational. - Ein Beispiel: Was ist eine gute wissenschaftliche Theorie? Typische wissenschaftstheoretische Antworten lauten: eine solche, die empirisch gehaltvoll ist, mit den Daten übereinstimmt, erfolgreiche Prognosen und Erklärungen liefert, usw..

Tatsächlich trifft dies auf die meisten Paradebeispiele anerkannter wissenschaftlicher Theorien zu, andererseits läßt es sich auf aufgrund des normativen Korrektivs rechtfertigen, daß zutreffende gehaltvolle Prognosen wohl das beste Indiz für das Vorliegen einer gehaltvollen wahren Theorie (Erkenntnisziel) sind. Dasselbe Verfahren wendet die Wissenschaftstheorie bei allen anderen elementaren Fragen an, z.B. was ist eine wissenschaftliche Erklärung?, oder wann ist eine Theorie als bewährt oder falsifiziert anzusehen?, etc.

Freilich muß die Zielsetzung der logischen Rekonstruktion, nämlich deskriptive Analyse und normative Rechtfertigung zusammenbringen, *nicht immer aufgehen*. Es ist nicht zu erwarten, daß alles an Wissenschaft faktisch Vorliegende auch rational gemäß dem normativen Korrektiv ist. In der Praxis verfährt die Wissenschaftstheorie, wie schon oben angeschnitten, so: sie bezieht sich zunächst nur auf *Musterbeispiele* wissenschaftlicher Erkenntnis - solche Beispiele von Theorien, die allgemein anerkannt sind und an deren Erfolg nicht gezweifelt wird. Die Wissenschaftstheorie versucht ihre Modelle zunächst mit diesen Musterbeispielen in Einklang zu bringen. Zugleich sollen die wissenschaftstheoretischen Modelle aber auch alle *Mustergegenbeispiele* ausschließen, also jene Musterbeispiele von Fällen, die unbezweifelbar als unwissenschaftlich anzusehen sind (*Abgrenzungsproblem*). Hat man ein wissenschaftstheoretisches Modell gefunden, das diesen Musterbeispielen+gegenbeispielen entspricht und zugleich dem normativen Korrektiv genügt, so wird man dieses Modell als adäquat oder bewährt ansehen. Nun kann man dieses Modell auch auf faktische wissenschaftliche Beispiele anwenden, die man als problematisch ansieht, wo also keine Einigkeit über deren wissenschaftliche Rationalität (Bewährtheit, Erklärungskraft, etc.) vorliegt. Und hier vermag die Wissenschaftstheorie dann den Einzelwissenschaften inhaltliche Entscheidungshilfen in die Hand zu geben. Einerseits also lernt die Wissenschaftstheorie aus den faktischen Wissenschaften, wird durch sie korrigiert; andererseits kann sie später, nach erfolgreicher Rekonstruktionsarbeit, den Wissenschaften helfen. Beispiele für beide Fälle werden wir noch kennenlernen. - Zusammengefaßt zeigt dies folgendes Schema:

NORMATIVES KORREKTIV

Allgemeines Erkenntnisziel

Allgemeines, minimales erkenntnistheoretisches Modell

(Voraussetzung)

WISSENSCHAFTSTHEORETISCHE REKONSTRUKTION

Modelle von Beobachtung, Experiment
Gesetz, Theorie, Erklärung, Bestätigung,
Falsifikation, Theorienfortschritt, usw.

(Analyse)

(Anwendung)

Faktische Wissenschaft
Musterbeispiele+ -gegenbeispiele
DESKRIPTIVES KORREKTIV

Faktische Wissenschaft
Problematische Fälle

Nicht immer aber läßt die Wissenschaftstheorie das erwähnte normative Korrektiv unbehelligt. Man kann vielmehr in der Wissenschaftstheorie in derselben Weise zwischen einer 'normalwissenschaftlichen' und einer 'revolutionären' Phase unterscheiden, wie dies von Kuhn für die Geschichte der Naturwissenschaften aufgezeigt worden ist. Genauer gesehen lassen sich sogar drei Phasen der Wissenschaftstheorie ausmachen:

1. *eine normalwissenschaftliche Phase*: in dieser zweifelt die Wissenschaftstheorie nicht an dem von ihr vorausgesetzten normativem Korrektiv, dem allgemeinen Erkenntnisziel und dem minimalen Modell wissenschaftlicher Rationalität.
2. *eine revolutionäre Phase methodologischer Art*: hier zweifelt die Wissenschaftstheorie an ihrem allgemeinen und minimalen erkenntnistheoretischen Modell, und zwar an den Annahmen (2)-(5) (während Annahme (1) bereits die Bedeutung von 'Wahrheit' und somit indirekt schon das Ziel betrifft.). Derartige Zweifel wurden beispielsweise durch die erwähnten Schriften von Thomas Kuhn (1976) und Paul Feyerabend (1976) ausgelöst. Diese Autoren versuchten unter anderem zu zeigen, daß alle Beobachtungen theoriebeladen sind und somit empirische Wissenschaft nicht 'objektiv' im erwähnten Sinne sein könne. Manche schlossen daraus, daß der Empirismus allein zuwenig sei und man einen starken Realismus bräuchte; andere, daß Wissenschaft überhaupt ein ziemlich irrationales, jedenfalls wenig objektives Unterfangen wäre, usw. In einer solchen Phase geht Wissenschaftstheorie direkt über in Erkenntnistheorie. Auch hier gibt es natürlich Möglichkeiten rationaler und nichtdogmatischer Begründung. Man kann z.B. mit Bezug auf die gesamte historische Entwicklung des menschlichen Denkens argumentieren, daß das logisch-empirisch-kritisch orientierte Denken die besten Erfolge erzielt hat, und andererseits die wenigsten Dogmen und Sekten hervorgebracht hat. Dennoch werden entscheidungsfähige Argumente in diesem Bereich wesentlich schwieriger.
3. *eine revolutionäre Phase normativer Art*, wo selbst das allgemeine Ziel der Wahrheitssuche angezweifelt wird bzw. die Bedeutung von 'Wahrheit' (Annahme 1)

neu definiert wird. So etwas wäre in der Tat ein kulturepochaler Umschwung, vergleichbar dem Übergang von der Antike zum Mittelalter oder vom Mittelalter zu neuzeitlicher Naturwissenschaft. Es ist klar, daß sich mit einer solchen Änderung auch die gesamte Wissenschaftstheorie ändern wird. Wie kann auf dieser Ebene noch rational argumentiert werden? Z.B. indem gezeigt wird, daß das allgemeine Erkenntnisziel der modernen Wissenschaft besser den erkenntnisexternen Lebensinteressen der Menschen dient als frühere historische Erkenntnisziele - Auf diese problematische Frage kann hier nicht näher eingegangen werden. Jedenfalls geht in der revolutionären Phase normativer Art die Wissenschaftstheorie direkt in normative Ethik über.

1.3 Das minimale gemeinsame erkenntnistheoretische Modell der Wissenschaft: fünf allgemeine Annahmen.

A1. Minimaler Realismus: Es gibt eine subjektunabhängig existierende Realität. Eine wissenschaftliche Disziplin sucht wahre inhaltvolle Aussagen über einen abgegrenzten Bereich derselben aufzustellen. Wahr (falsch) wird im Sinne der Korrespondenztheorie (s. oben) verstanden. Sätze sind umso inhaltvoller, je mehr Konsequenzen,- vornehmlich empirische Konsequenzen - sie besitzen. Somit ist klar, daß inhaltvolle Aussagen vorallem allgemeine Aussagen sein werden (s.u.). Auch andere Kriterien des Gehalts sind wichtig, z.B. *kausale Kontrolle* versus bloße *empirische Regularität*.

A2. Fallibilismus, kritische Einstellung: Es gibt keinen privilegierten und sicheren Weg zu Wahrheit in obigem Sinne; weder durch Intuition noch sonstige 'höhere Evidenzen'. Jede wissenschaftliche Behauptung ist - mehr oder minder - fehlbar; wir können uns ihrer Wahrheit nie ganz sicher sein. Daher kommt alles darauf an, durch empirische Überprüfung (s. 5) herauszufinden, ob eine wissenschaftliche Hypothese eher wahr ist, also z.B. wahrscheinlich oder bewährt ist, oder eher falsch ist, also z.B. geschwächt oder, im Extremfall, falsifiziert ist.

A3. Objektivität: Die Wahrheit einer Aussage muß aufgrund von Annahme 1 objektiv sein, d.h. unabhängig von den Einstellungen und Wertungen des Erkenntnissubjekts bestehen, da ja auch die Realität unabhängig davon besteht, und Wahrheit Übereinstimmung von Aussage und Realität ist. Diese Definition von Objektivität hilft uns in der Wissenschaftspraxis natürlich noch nicht weiter. Ein wichtiges (nicht sicheres, aber plausibles) Kriterium für Objektivität, und indirekt auch Wahrheit, ist jedoch *Intersubjektivität*: wenn die Wahrheit einer Aussage überhaupt sich überzeugend begründen läßt, so muß jeder kompetente Wissenschaftler oder Mensch von der Wahrheit einer Aussage überzeugbar sein. -- Man beachte, daß der Begriff der Objektivität neben der hier verwendeten Bedeutung als Subjektunabhängigkeit

noch in einer *anderen* Bedeutung eine Rolle spielt, nämlich in der Bedeutung von *vollständiger Wahrheit*. Dieser Bedeutung nach ist eine Sachverhaltsdarstellung objektiv, wenn sie ein möglichst vollständiges und ausgewogenes Bild aller relevanten Aspekte des Sachverhaltes liefert. Diese Bedeutung von Objektivität wird später, bei Kausalgesetzen und Theorien, wichtig.

A4. Logik im weiten Sinn: Durch Anwendung präziser logischer Techniken - für die Einführung und Definition von Begriffen, für die Formulierung von Sätzen sowie die Bildung korrekter Argumentationen - kann man am effektivsten dem Ziel der Wahrheitssuche (gemäß 1-3) näher kommen. Denn nur für präzise formulierte Sätze sind die logischen Konsequenzen *präzise ermittelbar*, und nur wenn letzteres der Fall ist, sind die Sätze präzise überprüfbar (gemäß. 5).

A5. Minimaler Empirismus: Der Gegenstandsbereich einer Wissenschaft - abgesehen von Formalwissenschaften wie Mathematik, auf die wir später zu sprechen kommen - muß im Prinzip der Erfahrung, der Beobachtung zugänglich sein. Denn nur durch Beobachtung kann verlässliche Information über die Realität erlangt werden. Beobachtungswissen erfüllt das Kriterium der Objektivität - es ist intersubjektiv, d.h. von subjektiven Voreinstellungen und Wertungen unabhängig. Empirische Beobachtungen sind also ein zentraler Schiedsrichter für wissenschaftliche Wahrheitssuche: an ihnen müssen wissenschaftliche Gesetze und Theorien empirisch überprüft werden. Beobachtungen sind zwar ebenfalls nicht infallibel (gemäß 2), aber bei ihnen ist Intersubjektivität am leichtesten und schnellsten erzielbar - Beobachtungssätze sind also epistemisch privilegiert.

Einige Erläuterungen: "Minimal" nennen wir den Realismus, weil er bloß die subjektunabhängige *Existenz* der Wirklichkeit behauptet - die Wirklichkeit ist also nicht etwa erträumt. Dies impliziert jedoch noch nicht, daß diese Wirklichkeit subjektunabhängig bzw. objektiv erkennbar ist - dies kann nur durch den 'Erfolg' der wissenschaftlichen Erkenntnismethode beantwortet werden. Wir behaupten lediglich in Punkt 4 und 5 zwei "minimale" objektive Ebenen, in denen Erkenntnis einen Prüfstein hat, der von subjektiven Voreinstellungen unabhängig ist: einerseits die Logik im weiten Sinn - die von sich aus jedoch noch nichts über die Wirklichkeit besagt, sondern nur universelles Denkwerkzeug ist; und andererseits die Beobachtung, die uns einen intersubjektiven und neutralen Zugang zur Wirklichkeit gibt. -- Der in Punkt 5 implizierte Empirismus ist minimal, weil keineswegs behauptet wird, daß sich alle wissenschaftlichen Behauptungen durch Definitionsketten auf Beobachtungen *zurückführen* lassen, oder gar durch sie beweisbar sind. Im Gegenteil sind, wie wir sehen werden, *die wenigsten wissenschaftlichen Sätze echte Beobachtungssätze*. Alles, was gefordert wird, ist, daß wissenschaftliche Sätze empirische Konsequenzen haben, an denen sie überprüfbar sind. Es ist durch unseren

'minimalen' Realismus und Empirismus nicht gesagt, inwieweit durch die wissenschaftliche Methode Objektivität und Wahrheit von Theorien tatsächlich erreicht werden kann, inwieweit der subjektive Spielraum von wissenschaftlichen Theorien eingeschränkt oder reduziert werden kann. Das kann - erneut - nicht apriori vorausgesetzt werden, sondern nur über den 'Erfolg' der wissenschaftlichen Methode beurteilt werden. Das allgemeine Erkenntnismodell macht, zusammengefaßt, so wenig Annahmen wie möglich, jedoch soviel wie - wenn auch zumeist unausgesprochen - nötig, um sinnvoll gegenstandsbezogene (i.e. nicht rein formal-mathematische) Wissenschaft betreiben zu können.

1.4 *Die minimale gemeinsame Methodologie der Wissenschaft: vier Grundbestandteile.*

Das allgemeine Erkenntnisziel plus die minimalen erkenntnistheoretischen Annahmen(1)-(5) sagen an sich noch recht wenig. Aus ihnen folgen aber sehr plausibel die methodologischen Grundbestandteile wissenschaftlicher Erkenntnis, mit denen wir es im folgenden zu tun haben werden:

Vorklärung: ein Beobachtungssatz heißt auch Basissatz. Einen Basissatz nennen wir *aktual*, wenn er auf einer tatsächlichen Beobachtung beruht. Nichtaktuale Basissätze sind z.B. empirische (singuläre) Voraussagen, die noch nicht beobachtet wurden.

M1. Wissenschaft sucht nach möglichst *allgemeinen und gehaltvollen Sätzen, Gesetzen und Theorien*, die in einer wissenschaftlichen Sprache abgefaßt sind (der Gesetzesbegriff wird so weit gefaßt, daß er sowohl deterministische wie statistische Gesetze, sowohl raumzeitlich unbeschränkte wie beschränkte umfaßt).

M2. Wissenschaft sucht nach *aktualen Basissätzen (Beobachtungssätzen)*, die die empirischen Beobachtungen bzw. Daten wiedergeben, und durch Beobachtung oder Experiment direkt bewahrheitet (bzw. bekannt) werden,

M3. Wissenschaft versucht, mithilfe der allgemeinen Sätze in 1. die bereits bekannten empirischen Tatsachen, ausgedrückt in aktualen Basissätzen gemäß 2., zu *erklären*, sowie noch unbekannte neue (z.B. in der Zukunft liegende) Basissätze *vorauszusagen*. (Hinweis: Von einer Erklärung spricht man dann, wenn das Erklärte auch tatsächlich der Fall ist; während dies bei einer Voraussage nicht vorausgesetzt wird). Erklärungen und Voraussagen haben oft deduktive Struktur, entsprechen also logischen Ableitungen; sie können aber auch induktiv-statistische Form haben.

-- Auf zweiter Ebene versucht die Wissenschaft auch, empirische Gesetze durch übergeordnete Theorien in diesem Sinn erklären oder vorauszusagen, welche ihrerseits im Erklärungsfall durch Basissätze gemäß 4 bestätigt wurden, im

Voraussagefall bestätigt werden sollen. Es gibt also drei Ebenen (s.u.).

M4. Wissenschaft versucht, ihre allgemeinen Gesetze und Theorien *empirisch zu überprüfen*, und zwar indem die vorausgesagten Basissätze mit den durch aktuellen Basissätzen verglichen werden. Stimmen die letzteren mit den ersten überein, so war die Voraussage erfolgreich (sie wird dann zu einer erfolgreichen Erklärung) und das Gesetz bzw. die Theorie ist *bewährt (bzw. bestätigt)*, stehen die letzteren mit den ersteren in Widerspruch, so war die Voraussage erfolglos und das Gesetz bzw. die Theorie ist falsifiziert (bzw. im statistischen Fall zumindest geschwächt).

Offenbar folgen diese Grundbestandteile 1-4 plausibel aus dem Erkenntnisziel und den minimalen erkenntnistheoretischen Annahmen 1-5: Gesetze und Theorien sind Musterbeispiele empirisch gehaltvoller Sätze - daher 1.; um etwas über deren Wahrheit auszusagen, müssen wir zunächst empirische Daten haben - daher 2. - mit denen wir die Sätze in 1. aufgrund ihrer empirischen Konsequenzen, d.h. aufgrund ihrer potentiellen Erklärungen und Voraussagen, empirisch überprüfen - daher Schritte 3. und 4.

Erfolgreiche Voraussagen und Erklärungen sind aber nicht nur für das wissenschaftliche Oberziel der "Findung wahrer gehaltvoller Sätze" wesentlich, sondern stellen zugleich die *zwei wichtigsten wissenschaftlichen Unterziele* dar (die dem Oberziel als 'Mittel' dienen, zugleich aber *darüber hinausgehen*): Voraussagen sind für das technische und praktische Handeln und Planen eminent wichtig und Erklärungen für das intellektuelle Bedürfnis, die Welt zu verstehen.

Zusammengefaßt gibt es drei Ebenen:

Ebene übergeordneter Theorien

Erklärung/Voraussage Bewährung/Schwächung

Ebene empirischer Gesetze

Erklärung/Voraussage Bewährung/Schwächung

Ebene von Basissätzen

Ein Beispiel soll dies unsere 4 Merkmale und 3 Ebenen verdeutlichen:

THEORIE:

1. CO im Rauch reduziert O₂ im Blut + weiteres Hintergrundwissen
2. O₂ - Mangel führt zu Schwindelgefühl
 - Theoriemodifizierung: 2.' ... aber auch O₂-Überschuss führt zu Schwindelgefühl

Voraussagen

Erklärung

Neue Gesetzesphänomene:

Schwindel bei
HyperatmungSchwindel bei
starkem Smogbereits bekanntes
GesetzesphänomenEMPIRISCHES
GESETZ

Rauchen führt zu Schwindelgefühl

Gesetzesmodifizierung: ... außer bei
Gewohnheitsrauchern

Erklärung

Voraussage

Einige bereits
bekannte Fälle

Neue Fälle:

z.B. "mir wurde
nach dem
Rauchen schwindlig"ein
Gewohnheitsraucherein
Nichtraucher

Bestätigung:

Schwächung:

1.5 Drei Arten der Induktion

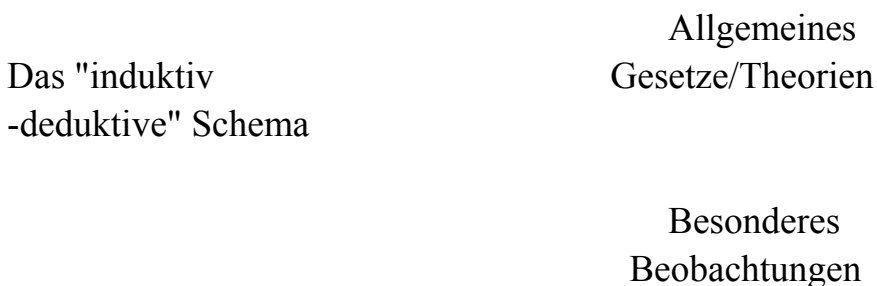
In spielt nicht nur logisches bzw. deduktives Schließen (z.B. von Gesetz und Anfangsbedingung auf Prognose) eine Rolle, sondern vorallem auch *induktives Schließen*. Darunter versteht man ganz allgemein den Schluß von bisherigen Beobachtungen auf einen allgemeinen Gesetzeszusammenhang (bzw. auch, auf neue bisher unbeobachtete Fälle -- Voraussagen). Im Gegensatz zu logischen Schlüssen sind induktive Schlüsse nicht "sicher" - das heißt, gegeben ihre Prämissen sind wahr, so ist die Konklusion nicht "sicher" wahr, so wie bei logischen Schlüssen. Man sagt

auch, induktive Schlüsse sind *gehaltserweiternd*, und daher unsicher, während logische Schlüsse *gehaltsbewahrend* und daher sicher sind.

Es war David Hume, der als erster gezeigt hat, daß induktive Schlüsse nicht mit logischen Mitteln begründet werden können. Die Rechtfertigung induktives Schließens ist an anhaltendes wissenschaftstheoretisches Problem, das wir hier nicht besprechen. Im folgenden sollen nur die drei Hauptarten besprochen werden, in denen "Induktion" im Laufe der Geschichte der Wissenschaftstheorie verstanden wurde.

1.5.1 Die methodische Induktion

In dieser Auffassung ist Induktion eine Methode, um von Besonderen (dem Einzelfall, der Beobachtung) auf das Allgemeine (das Gesetz, die Theorie) zu schließen, während Deduktion eine Methode darstellt, um von Allgemeinen zurück auf das Besondere zu schließen (deduktive Voraussage, Erklärung). Dieses "induktiv-deduktive" Schema findet sich letztlich bereits bei Aristoteles.



Obwohl vieles an dem Schema in starker Vereinfachung stimmt, ist das Schema doch auch in mehreren Hinsichten inkorrekt. Zu einem stimmt es nicht, daß die Induktion immer vom Besonderen zum Allgemeinen schießt, es gibt auch sogenannte induktive Spezialisierungsschlüsse, aber damit beschäftigen wir uns nicht. Der Hauptfehler dieses Schemas liegt darin, wie vorallem Popper argumentiert hat, daß hier der Entdeckungszusammenhang mit dem Begründungszusammenhang verwechselt wird. In diesem Schema ist Induktion vorallem eine Methode, um allgemeine Gesetze oder Theorien aus besonderen Beobachtungen zu gewinnen, sie zu entdecken. Aber, wie Popper argumentiert, für die wissenschaftliche Begründung einer Theorie ist es irrelevant, *wie* diese entdeckt worden ist. Im Fall einfacher Gesetzhypothesen wie "Alle Raben sind schwarz" mag eine solche Entdeckung tatsächlich durch simple induktive Verallgemeinerung der bisher beobachteten schwarzen Raben erfolgt sein. In vielen Fällen komplexerer z.B. physikalischer Theorien kann eine solche Entdeckung ganz anderes erfolgen, z.B. durch theoretische Spekulation, durch Analogiebildung, oder schlicht durch Einfall (wir erwähnten bereits Kekule, der den

Benzolring erträumte). Wichtig ist dann nur, wie die - irgendwie entdeckte -. theoretische Hypothese - später wissenschaftlich begründet wird- durch empirische Überprüfungen im Sinn unserer methodischen Merkmale M1-M4.

1.5.2 Die logische Induktion

Induktivisten wie Carnap und Reichenbach haben diesen Punkt Popper längst zugestanden. Sie verstanden Induktion als eine Methode der Begründung von allgemeinen Hypothesen bei gegebenen Beobachtungssätzen.

Die logische Induktionsauffassung geht auf Carnap zurück und besteht in dem Versuch, ein System induktiver Prinzipien zu errichten, innerhalb derer man die Wahrscheinlichkeit (die Glaubenswahrscheinlichkeit) einer theoretischen Hypothese H, gegeben die und die aktuellen Beobachtungen B, verlässlich feststellen bzw. ermitteln kann. Die logische Induktion meint also, man könne zu Bewertungen der folgenden Form gelangen:

Wahrscheinlichkeit(Hypothese H / Beobachtung B) = so und so (z.B. 0,83)

Auch hier hat Popper die entscheidende Kritik geübt. Das Problem dabei ist, daß Aussagen über den Absolutwert der Wahrscheinlichkeit einer Hypothese bei gegebener Beobachtung nur möglich ist, wenn es nur endlich viele Alternativtheorien gibt, und man alle diese kennt. Ein Beispiel wäre die Hypothese "70% aller Schweden sind blondhaarig" im Raum der Alternativhypothesen "x % aller Schweden sind blondhaarig", wobei x alle Werte von 0 bis 100 durchläuft. Das Wesentliche wissenschaftlicher Theorien ist es jedoch, daß diese neue, sogenannte theoretische Begriffe enthalten, die über die Beobachtungsbegriffe hinausgehen. Die Möglichkeit der "Erfindung" neuer theoretischer Begriffe und damit neuer alternativer Theorien ist daher logisch nicht begrenzt. Aus diesem Grund ist die Feststellung von absoluten Werten für die Wahrscheinlichkeit wissenschaftlicher Theorien (bei gegebenen Daten) nicht möglich.

1.5.3 Die pragmatische Induktion

Diese Induktionsart ist in die zuvor eingeführte Methode der Theorienüberprüfung bzw. Theorienbewährung eingebaut und besteht darin, den *bisher* gut bewährten Theorien auch in Zukunft eher zu vertrauen. Der pragmatische Induktionsschluß hat folgende Form:

T₁ hat sich bisher besser

T₁ wird auch in Zukunft

bewährt als T_2

mehr vertraut als T_2

Wichtig ist, daß diese pragmatische Induktion immer nur komparativ ist - d.h. es wird der Bewährungsgrad zweier oder mehrerer Alternativtheorien miteinander verglichen (aufgrund der bisherigen Bewährung). Es werden aber keine Absolutaussagen über die Bewährtheit oder die "Wahrheitsnähe" von Theorien aufgestellt. Auch wenn eine gegebene Theorie T bisher viel besser bewährt war als aller bisher verfügbaren Alternativtheorien (z.B. die Newtonische Physik vor Einstein), so kann doch schon morgen jemand eine neue Theorie T^* "erfinden", die sich im weiteren Verlaufe dann als der Theorie T weit überlegen herausstellt, und die damit indirekt demonstriert, daß T dennoch ziemlich weit von der Wahrheit entfernt war.

Gemäß der pragmatischen Induktion ist der Bewährungsgrad einer Theorie *doppelt relativ* -

- relativ zum gegebenen Stand des Beobachtungswissens, sowie
- relativ zum derzeitigen Stand der Alternativtheorien.

Wie gesagt, ist die pragmatische Induktion implizit im Bewährungsbegriff eingebaut. Popper hat zwar zeitlebens jegliche Art der Induktion verurteilt, bei genauerer Analyse zeigt sich, daß auch er in seinem Bewährungskonzept von einem solchen pragmatischen Induktionsprinzip Gebrauch macht - es gibt auch Stellen (etwa in "Objektive Erkenntnis"), wo er dies selbst sagt.

1.6 Untergliederung wissenschaftlicher Disziplinen und Abgrenzung von "Wissenschaft"

Wir wollen nun untersuchen, inwieweit unsere minimalen erkenntnistheoretischen Annahmen 1-5, sowie die Grundbestandteile der wissenschaftlicher Methode 1-4 auf die existierenden Wissenschaftsdisziplinen zutreffen. Zugleich werden wir nach einer plausiblen und undogmatischen "Abgrenzung" des Begriffs "Wissenschaft" suchen - wobei hierfür insbesondere die Annahmen 1-5 entscheidend sein werden, während die Bestandteile 1-4 durchaus in verschiedenen Disziplinen variieren in verschiedenem Ausprägungsgrad vorliegen können. Gehen wir von folgender Einteilung von Disziplinen - mit lediglich heuristischem Wert - aus (nicht explizit genannte Spezialdisziplinen sind entsprechend einzufügen):

1) Wissenschaften von der Natur:

Physik, Chemie, Biologie, Geologie, Ökologie, und alle ihre technischen Anwendungsdisziplinen, inklusive Medizin.

2) Wissenschaften vom Menschen:

Psychologie (Pädagogik)

3) *Wissenschaften von der Gesellschaft:*

Ökonomie, Soziologie, Politikwissenschaft

4) *Wissenschaften vom Menschen und der Gesellschaft in der Geschichte:*

Geschichte, Anthropologie, Ethnologie

5) *Wissenschaften von geistigen (sozialen und kulturellen) Schaffensprodukten des Menschen:*

Sprachwissenschaft (Publizistik), Rechtswissenschaft, Literaturwissenschaft, Kunstwissenschaft (Pädagogik)

6) Struktur-, Methoden- und Metawissenschaften

spezielle: Mathematik (inklusive mathematische Logik), Statistik (Informatik, Systemtheorie, Kybernetik,...), Angewandte Logik verschiedenster Bereiche, Spezielle Wissenschaftstheorien

allgemeine:

Philosophische Logik, Allgemeine Wissenschaftstheorie, Philosophie der Wissenschaften

7) *Wissenschaften von allgemeinen Grundlagen geistiger und sozialer Schaffensprodukte des Menschen:*

Philosophie:

Grundlagen der Erkenntnis: Erkenntnistheorie, Ontologie, Philosophien einzelner Bereiche (Natur-, Sozial-, Sprachphilosophie, usw.).

Grundlagen der Moral: Ethik

Grundlagen des Schönen: Ästhetik

8) Theologie und Religionswissenschaft, Religionsphilosophie.

Folgende Bezeichnungen sind auch üblich: Bereich (1) - die Naturwissenschaften, nach mehrheitlichem Wunsch der Psychologen heute auch Bereiche (1)-(2); Bereich (2)-(3) auch Human- und Sozialwissenschaften, mit Sonderstellung der Ökonomie als Wirtschaftswissenschaft; Bereich (1)-(5) ist der Bereich der Substanzwissenschaften i.w.S., im Gegensatz zum Bereich (6) (und Teile von (7), s.u.) als Struktur- und Methodenwissenschaften); Bereich (4)-(5), (7) und (8) heißt auch Geisteswissenschaften. Der Unterschied zwischen (1)-(7) und (8) liegt darin, daß (1)-(7) vom Irdischen, (8) vom Überirdischen handeln.

Der Abkürzung halber nennen wir im folgenden die minimalen fünf erkenntnistheoretischen Annahmen A1-A5; die minimalen methodologischen Grundbestandteile M1-M4. Es ist offensichtlich, daß A1-A5 sowie M1-M4 auf (1), die Naturwissenschaften, paßt. Es spricht auch nichts Zwingendes gegen die Anwendbarkeit auf Bereiche (2) und (3). Hier gibt es überall empirische Daten, und zumindest im Prinzip exaktifizierbare wissenschaftliche Sprachen; man sucht nach

allgemeinen Gesetzen und Theorien, die man empirisch zu überprüfen sucht. Natürlich sind in Bereichen wie Psychologie und Soziologie allgemeine Gesetze und Theorien wesentlich schwerer auffindbar - doch lassen sich hier statistische Gesetze ausfindig machen. Ebenfalls gibt es im Bereich (4), den Geschichtswissenschaften, eine empirische Datenbasis, eine wissenschaftliche Sprache und dgl. Allerdings wurde oft betont - aus unterschiedlichen Gründen - daß es in der Geschichte keine allgemeinen Gesetze gäbe (z.B. Popper, Wilhelm Dilthey, William Dray). Diese Einwände betreffen eher die Merkmale M1-M4 als die Annahmen A1-A5. Es gute Argumente dafür, daß es in der Geschichte zumindest statistische Trendgesetze gibt. Wie dem auch sei, so ist dies doch nur ein gradueller Unterschied hinsichtlich des Bestandteils M1. Und wohlgemerkt ist es für M1-M4 nur wichtig, daß die jeweilige Wissenschaft nach den darin genannten methodologischen Zielen sucht, was nicht schon heißt, daß sie es überall gleichgut erreichen wird können. Natürlich gibt es in der Physik zweifellos viel mehr Gesetze als etwa in der Geschichtswissenschaft, weil in letzterer Gesetze ungemein schwer aufzufinden sind; wesentlich für das Zutreffen von M1 auf die Geschichtswissenschaft ist jedoch lediglich, daß sie - unter anderem - auch danach sucht. Ferner gibt es in Psychologie wie Soziologie bekannte methodologische Schulenstreite; doch daß man angesichts der heute existierenden empirischen Psychologie und Soziologie diese Disziplinen als empirische Wissenschaft betreiben kann, scheint doch unbezweifelbar.

Anders aber wird es schon in Disziplinen vom Bereich (5). Die Sprachwissenschaften, die die Struktur und historische Entwicklung von Sprachen studieren, scheinen noch relativ unproblematisch in unser Modell einzuordnen sein. Gehen wir nur aber zu den *Rechtswissenschaften*. Sofern in der reinen Rechtstheorie, -historie - und -soziologie (usw.) empirisch-historisch vorliegende Rechtssysteme bloß beschrieben und erklärt werden, ist diese Tätigkeit im Einklang mit den Annahmen A1-A5 (sowie, in unterschiedlicher Ausprägung, ebenfalls mit M1-M4). Was in der juristischen Dogmatik und Praxis jedoch gemacht wird, ist folgendes: Es werden konkrete empirische Tatbestände daraufhin untersucht, wie sie sich in das gegebene (gegenwärtige) Rechtssystem einordnen lassen - unter welchen Paragraphen sie fallen. Z.B.: war eine bestimmte Entwendung von Privateigentum *Mundraub* oder *Diebstahl* ? Aus den Definitionen des Gesetzbuchs und der detaillierten empirischen Tatschilderung folgt jedoch in der Regel noch nicht, worunter die konkrete Tat fällt; und der Richter oder praktisch arbeitende Rechtswissenschaftler hat hier die Aufgabe, die Gesetzesbestimmungen *auszulegen*, d.h. durch zusätzliche Bedeutungspräzisierungen der im Gesetzestext stehenden Begriffe soweit zu ergänzen, bis aus Gesetzestext, zusätzlichen Auslegungsannahmen und konkreter Tatbeschreibung logisch folgt, daß die Tat unter einen bestimmten Paragraphen fällt. Wie Eike von Savigny (1976) nachwies, stützt man sich dabei oft

auf *vorausgesetzte ethische Annahmen*, die nicht im Gesetzestext stehen, sondern dem *intuiven Rechtsbewußtsein des Normalbürgers* entnommen werden (man nennt solche ethische Zusatzannahmen auch Billigkeitserwägungen). Nun kann der Rechtswissenschaftler sicher zeigen, daß unter Voraussetzung gewisser zusätzlicher ethischer Annahmen die und die Auslegung adäquat ist. Doch kann er diese ethischen Annahmen selber im wissenschaftlichen Sinne begründen oder sie in irgendeinem wissenschaftlichen Sinne überprüfen? Das ist die entscheidende Frage, die wir nun kurz diskutieren wollen.

In medias res - die von uns vorgeschlagene Antwort darauf lautet: "nein", und dies soll uns zugleich dazu dienen, den wissenschaftlichen Bereich vom nichtwissenschaftlichen abgrenzen. Die Auffassung, daß Wissenschaft keine Werte - in fundamentalen Sinn - begründen kann, steht in der Tradition des herkömmlichen Wissenschaftsverständnisses; am brilliantesten wird sie wohl in Max Webers "Wissenschaftslehre" vertreten und begründet. Was sind die Hauptgründe für diese "Selbstbeschränkung" der Wissenschaft? Nun, offenbar können moralische Normen wie "Du sollst den Nächsten lieben", oder Werte wie "Freiheit", "Gerechtigkeit", "Wohlstand" nicht so wie empirische Fakten beobachtet oder durch empirische Verallgemeinerungen gewonnen werden. Es sind nicht Dinge, die uns die Realität über unsere Wahrnehmungsorgane 'aufzwingt'. Es scheint uns bei der Frage der Überprüfung moralischer Sätze nichts zu geben, das der empirischen Basis gemäß A5 und M2 entspricht. Es liegt letztlich in der *Freiheit* des Menschen, sich zu gewissen Normen und Werten zu bekennen. Daher gibt es, unserer Auffassung nach, im Bereich der Normen und Werte keine Objektivität im wissenschaftlichen Sinne. Somit scheint uns hier, im Bereich der Norm- und Wertebegründung - man spricht auch von normativer Ethik - nicht mehr von Wissenschaft gesprochen werden zu können.

Das heißt nun nicht, daß die Rechtswissenschaft keine Wissenschaft gemäß unserem Modell ist - die daraus folgende Aussage ist differenzierter: Soweit die Rechtswissenschaft entweder empirisch vorliegende Rechtssysteme bloß analysiert, oder bei der Auslegung und Deutung von Rechtssystemen zum Zweck der Einordnung empirischer Tathandlungen sich auf unproblematisch gegebene normative Hintergrundannahmen stützt und logische Argumentationsketten konstruiert, ist sie Wissenschaft. Mit anderen Worten, der Schluß von einer fundamentalen Norm und einer deskriptiven Gesetzesprämisse auf eine abgeleitete Norm, im Sinn des Zweck-Mittel-Schemas, ist sehr wohl eine wissenschaftsimmanente Tätigkeit, und kann vom Wissenschaftler qua Wissenschaftler behauptet werden, sofern dabei die ethische Fundamentalnorm, auf die man sich stützt, nur nicht *verschwiegen* wird, sondern explizit als *ethische Voraussetzung* des Erschlossenen benannt wird. Der Zweck-Mittel-Schluß hat dabei

im einfachsten Fall folgende Form:

A ist ein (fundamentaler) Wert

A läßt sich nur realisieren, wenn B getan wird

Daher ist B ein (abgeleiteter) Wert.

Sofern der Wissenschaftler aber ethische Prämissen im *fundamentalen Sinn* - also ohne Rekurs auf andere, vorausgesetzte ethische Prämissen - begründen will, verläßt er den Bereich der Wissenschaft.

Die logische Explikation und Systematisierung von ethischen Werten und Normen - insbesondere Zurückführen vieler auf einige wenige - betreibt nicht nur die juristische Dogmatik im oben erwähnten Fall, sondern vorallem die 'analytische Ethik' bzw. die Werte- und Normenlogik oder *Metaethik*. Wenn wir sagten, dieses beschränkte Unterfangen sei eine Art Wissenschaft, so muß hinzugefügt werden, daß diese durchaus nicht dieselbe methodologische Struktur wie die durch M1-M4 wiedergebene Struktur empirischer Wissenschaft: hier fehlt ja die 'Basis'. Es handelt sich hier um eine Tätigkeit, die grundsätzlich in den Bereich der Formal- und Metawissenschaften, oder vielleicht treffender Struktur- und Rekonstruktionswissenschaften, einzuordnen ist, auf dessen Sonderstellung wir noch zurückkommen werden. Kurz gesagt liegt die Sonderstellung dieser Wissenschaften darin, daß ihre Hauptaufgabe in der Explikation und Differenzierung von Begriffssystemen und in der logischen Konstruktion und Rekonstruktion von Argumentationen besteht - also im Grunde in einer logisch-begrifflichen Tätigkeit. Solche Unterfangen sind nicht selbst 'Substanzwissenschaften', aber unerläßliche Voraussetzungen für Substanzwissenschaften, die sich aber einem gewissen wissenschaftlichen Entwicklungsstand als selbständige Disziplinen ausdifferenzieren. In diesem Sinn kann auch Ethik eine Wissenschaft sein. Aber wesentlich ist, daß Ethik als *logische Rekonstruktionswissenschaft* ethische Werte im fundamentalen Sinn - also analog zur empirischen Wissenschaft durch Rekurs auf eine objektive Basis - weder begründen kann noch begründen will. Kurz gesagt, Ethik kann unserer Auffassung nach keine Substanzwissenschaft sein.

In dieser Linie können wir nun alle weiteren noch ungenannten Disziplinen analysieren. Beispielsweise beschäftigt sich die Literaturwissenschaft und Kunstwissenschaft großteils mit empirisch vorliegendem Literatur- und Kunstmaterial eines *Stilbereichs*, und entwickelt verschiedene Theorien über die Merkmale, die für die *ästhetische Wirkung* auf den mit dem Stil vertrauten Menschen verantwortlich sind. Soweit hier unter ästhetischer Wirkung eine faktische Wirkung auf den Menschen gemeint ist, und soweit ein abgegrenzter Bereich ästhetischer Objekte untersucht wird, passen diese Disziplinen durchaus in das Schema

empirischer Wissenschaft. Aber sofern hier allgemeine Aussagen über Schönheit als *Wert* gemacht werden; über das, was an sich den ästhetischen Wert besitzt - ohne daß Schönheit mit empirischen Stilkriterien identifiziert wird, und ohne daß im Sinne einer Rekonstruktionswissenschaft bloß ästhetische Wertaussagen logisch auf andere zurückgeführt werden - kann nicht mehr von Wissenschaft gesprochen werden, da es hier um Begründung von Werten geht. Dasselbe gilt auch für die unter (2) erwähnte Politikwissenschaft und die unter (6) erwähnte Pädagogik, sobald dort politische Werte respektive Erziehungswerte begründet werden sollen. Was schließlich die Philosophie betrifft, so kann, wie schon erwähnt, jener Teil der Ethik, der Werte (im fundamentalen Sinn) begründet, nicht mehr Wissenschaft genannt werden. Die Religionswissenschaften schließlich stützen sich - sofern sie über reine empirische Religionsanalyse hinausgehen - nicht bloß auf wissenschaftlich unüberprüfbare Werte und Normen, sondern zudem noch auf unüberprüfbare und schlicht geglaubte deskriptive Behauptungen, die das "Credo" der Religion ausmachen. Als Substanzwissenschaften kann es sie daher noch weniger geben als normative Ethik. Selbstverständlich aber kann es Religionswissenschaft als empirisch-historische Wissenschaft oder 'Religionslogik' als Struktur- und Rekonstruktionswissenschaft geben.

Wir kommen nun zu (6), den Struktur-, Methoden- und Metawissenschaften. Wir haben schon einiges über deren *Sonderstellung* gesagt. Auch die Mathematik hat z.B. keine empirische Basis, d.h. ein Großteil der Annahmen A1-A5 und Bestandteile M1-M4 treffen auf sie nicht zu. Dennoch liegt die Situation hier ganz anders. Die Mathematik macht nämlich keine inhaltlichen, gehaltvollen Aussagen über die *Realität*, und auch nicht über die Wertsphäre. Sie ist eine reine *Formalwissenschaft*. Sie entwickelt mathematische Modelle und leitet daraus logische Konsequenzen ab. Die mathematischen Modelle sind unmittelbar überhaupt keiner Überprüfung ausgesetzt, weil sie nicht inhaltlich-empirisch interpretiert sind. Es ist sinnlos, zu sagen, das mathematische System der euklidischen Geometrie sei 'wahr' oder 'falsch', ebenso wie es sinnlos wäre, zu sagen, das der gekrümmten (Riemannschen) Geometrie sei 'wahr' oder 'falsch'. Solche Behauptungen kann man erst machen, wenn die Geometrien empirisch interpretiert sind, wenn den Vektoren z.B. reale Längen zugeordnet werden - dann ist die Geometrie aber nicht mehr um eine mathematische, sondern eine physikalische Theorie. Die Mathematik entwickelt also bloß formale Strukturen, die dann in den Substanzwissenschaften zur Anwendung gelangen. Somit ist die Mathematik eine Wissenschaft, die eine *Sonderstellung* unter den anderen Wissenschaften besitzt: sie ist eine *Voraussetzungs- und Hilfswissenschaft* für alle anderen Disziplinen, die formale Strukturen liefert. Aus diesem Grund treffen auf die Mathematik nur die Annahmen A1-A4 und der Bestandteil M1 unseres allgemeinen Wissenschaftsbegriffs zu - alles, was mit empirischer Überprüfung zu tun hat, fehlt.

In der selben Linie können wir auch die anderen in (7) genannten Disziplinen als Hilfswissenschaften auffassen: als Lieferanten von Methoden und Argumenten.

Den Bereich der Wissenschaften können wir zusammenfassend damit so abgrenzen:

Wissenschaften im hier verwendeten Sinn sind alle im weiten Sinne empirischen Disziplinen (gemäß Annahmen A1-A5) inklusive ihrer Hilfswissenschaften, also Methoden- und Metawissenschaften.

Wissenschaftspolitische Ergänzung: Es muß an dieser Stelle betont werden, daß unsere Abgrenzung von Wissenschaft *keine Wertung* impliziert, und schon gar nichts für die Frage, was an den Universitäten gelehrt werden soll. Unsere Wissenschaftskonzeption wurde aus wissenschaftstheoretischen Gründen getroffen: sie soll die *charakteristischen Merkmale eines homogenen historischen Erkenntnistyps namens "Wissenschaft" enthalten* und ihn dadurch auch gut erklären können; und es scheint uns aus Gründen der Begriffsklarheit nicht zweckmäßig, sich stark davon unterscheidende Geistestätigkeiten, wie Ethik oder Religion (im 'Fundamentalsinn'), mit demselben Namen "Wissenschaft" zu bezeichnen. Selbstverständlich aber sind auch andere Erkenntnisgebiete wie *Moral, Religion* usw. für das Geistes- und Kulturleben bedeutsam. Darüberhinaus gibt es in der Philosophie nicht nur die hier vertretene Wissenschaftskonzeption, sondern auch noch andere Konzeptionen, auf die wir später in einem Exkurs eingehen werden - etwa die insbesondere in Teilen der Literatur- und Kunstwissenschaft und der spekulativen Philosophie vertretene 'hermeneutische' (dialektische, spekulative) Wissenschaftskonzeption, die Annahme A4 unseres allgemeinen Erkenntnismodells bestreiten würde, d.h. nicht die präzisen, sondern eher spekulativen oder gar literarischen Sprachen bevorzugt; ferner die 'konstruktivistische' Wissenschaftskonzeption, die - jedoch nur in ihrer ganz radikalen Varianten - Annahme A1 nicht teilen möchte, usw.

Daß fundamentale Wertsätze nicht empirisch-wissenschaftlich begründbar sind, ist die mehrheitliche Auffassung der Wissenschaftstheoretiker. Es wird allerdings auch die Gegenposition vertreten, z.B. von Weingartner (1978). Dieser Gegenposition zufolge können auch Wertaussagen - ähnlich wie empirische Gesetzhypothesen - wissenschaftlich begründet werden. Wir besprechen kurz die wichtigsten Argumente dieser Gegenposition und zeigen, warum sie der hier vertretenen Auffassung zufolge nicht stichhaltig sind.

Das erste Argument besteht darin, daß die *Intersubjektivität* - welche wir ja als wichtigstes Merkmal von Beobachtungssätzen ansahen - etwas *Graduelles* sei, und nicht dazu hinreicht, einen grundlegenden Unterschied zwischen empirischen

Basissätzen und normativen (bzw. werthafte) Basissätzen zu machen. Vollkommene (100%ige) Intersubjektivität gibt es nämlich auch bei reinen Beobachtungssätzen nicht. Sehr hohe (z.B. 95%ige) Intersubjektivität gibt es aber auch bei Basisnormen wie "Peter soll seine Mutter lieben". Wir tragen folgendes Gegenargument vor:

Die Tatsache, daß einige Werte und Normen von einer großen Mehrzahl aller gegenwärtig lebenden Personen geteilt wird, bedeutet noch nicht Intersubjektivität im Sinn unseres Modells A3. Wir sagten nämlich: die Wahrheit eines Satzes S ist intersubjektiv, wenn jeder Mensch (und nicht fast jeder, oder mehr als 80%) unter gewissen '*Normalbedingungen*', die seine urteilsrelevanten Kompetenzen betreffen, der Wahrheit zustimmen *würde*. Das scheint uns auf Beobachtungssätze zuzutreffen. Um es gleich zu gestehen: die Schwierigkeit besteht in der richtigen Explikation des Normalbedingungen der Beobachtung. (Das ist jedoch nichts ungewöhnlich; überall, wo man in der Wissenschaft nach regelmäßigen Zusammenhängen sucht, muß man diese Schwierigkeiten lösen).- Betrachten wir ein Beispiel: zu beurteilen ist der angenommen wahre Beobachtungssatz "*dort brennt ein Feuer*". Die Normalbedingungen der Beobachtung beinhalten: die Person, die den Satz beurteilen soll, hat keinen (empirisch feststellbaren) Defekt der Sinnesorgane oder des Gehirns, sie kann den Satz "dort brennt Feuer" in ihre Muttersprache übersetzen, sie ist im psychischem Normalzustand (d.h. halluziniert nicht), ist wahrhaftig, und sie ist nahe genug am Ort des zu beobachtenden Ereignisses, um es sehen zu können (die Bedingung, daß Licht vorhanden ist, ist im Fall eines Feuers automatisch erfüllt). Wir behaupten: jede Person unter diesen Normalbedingungen würde der Behauptung "dort brennt ein Feuer" zustimmen. Würde eine Person das - hartnäckig und wiederholt - nicht tun, so hätten wir allen Grund, irgendeinen *Defekt in den Normalbedingungen* anzunehmen, und in allen solchen Fällen wäre wohl ein solcher Defekt nachweisbar. Wenn z.B. ein Mensch mit offenen Augen ins Feuer rennt, so ist er höchstwahrscheinlich blind oder halluziniert extrem. Und es wäre völlig berechtigt, so einen Menschen von seinem Beobachtungsdefekt zu heilen versuchen.

Doch nehmen wir an, unter denselben Normalbedingungen würde ein Erwachsener einer Basisnorm wie z.B. "daß eine Person sich das Leben nimmt, ist schlecht" systematisch nicht zustimmen (obwohl, angenommen, 99% aller Personen zustimmen). Dann wären wir nicht berechtigt anzunehmen, daß ein Defekt in den '*moralischen Normalbedingungen*' dieser Person vorliegt, d.h. daß diese Person z.B. *moralisch minderwertig sei und davon geheilt werden sollte*, genauso wie ein Blinder biologisch minderwertig ist und von seiner Blindheit, wenn möglich, geheilt werden sollte. Vielmehr müssen wir dies als *abweichenden Werthaltung* achten und - sofern dadurch nicht Interessen anderer geschädigt werden - auch zulassen. (Diese Auffassungsunterschiede haben offenbar weitere Konsequenzen, etwa bzgl. der Frage, ob die Findung ethischer Wahrheiten wissenschaftlichen Expertengruppen

überlassen werden darf- was gemäß unserer Auffassung jedenfalls nicht der Fall ist; außer, die Delegation geschieht freiwillig und aus organisatorischen Gründen).

Ein zweites Argument besagt, daß natürlich die Frage, wie wir unsere Beobachtungen moralisch interpretieren, von unseren ethischen Hintergrundannahmen abhängig sein -- aber im selben Sinn seien ja auch die empirischen Beobachtungen selbst von unseren deskriptiven Hintergrundtheorien abhängig (so haben ja Kuhn und Feyerabend argumentiert). Ein wichtiges Gegenargument hierzu stammt von *Toulmin*. Unser Beobachtungserlebnis wird nämlich in der Tat nicht durch unsere explizit-bewußten Hintergrundtheorien verändert. Beispielsweise sehen wir tagtäglich die Sonne über das Himmelsgewölbe wandern. Dieses Seherlebnis wird in keiner Weise dadurch beeinflusst, ob wir nun Heliozentriker sind und kognitiv glauben, die Sonne drehe sich um die Erde, oder ob wir Geozentriker sind, und kognitiv glauben, in Wirklichkeit drehe sich die Erde um die Sonne. Das Seherlebnis bleibt gleich. -- Ganz anders im moralischen Empfindungserlebnis. Es ändert sich durch anderes Hintergrundwissen. Sehe ich einen Menschen, der eine Million den Armen spendet, so hab ich die moralische Empfindung, das war eine gute Tat. Erfahre ich hinterher, daß diese Person ein Politiker kurz vor seiner Wahl war, der dies nur zum Zwecke der Vergrößerung von Wählerstimmen tat und sich ansonsten mehr um die Reichen als um die Armen kümmert, so werde ich angesichts dieser Tat eine ganz andere moralische Empfindung haben. (Beispiel ist von Toulmin). Hier zwingt sich also ein zweiter Unterschied: unsere moralische Empfindungen sind *im Gegensatz* zu unseren Wahrnehmungsempfindungen tatsächlich von unseren bewußten Hintergrundtheorien abhängig. Es handelt sich nicht um echte "beobachtungen", sondern eben um moralische *Interpretationen* von Beobachtungen.

Abschließend erwähnen wir, daß es noch andere Arten gibt, Wissenschaftsdisziplinen einzuteilen. So kann man die *Substanzwissenschaften* einteilen in:

- 1) "*spekulative*" Disziplinen (nichtempirisch aber synthetisch, nichtwissenschaftlich in unserem Sinn)
- 2) *empirische Wissenschaften* - alle die sich auf empirische Daten irgendwelcher Art beziehen, also auch z.B. Geschichtswissenschaften (historisches Quellenmaterial), Literaturwissenschaften (Texte), usw.
- 3) *experimentelle Wissenschaften* - solche, die nicht nur empirische Beobachtungsdaten besitzen, sondern diese in Form kontrollierter Experimente ermitteln. Auf die Bedeutung des Experiments gegenüber freier "Feldbeobachtung" kommen wir in Kap. 3 zurück. Diese Kategorie schließt viele "Geisteswissenschaften", etwa Geschichtswissenschaften, Literaturwissenschaften,

Makrosoziologie, etc. aus, weil man hier nicht experimentieren kann. Schließt aber z.B. die experimentelle Psychologie ein, und natürlich auch die klassischen Naturwissenschaften.

4) "*Sezierende*" *Wissenschaften* - das sind die, die ihre empirischen Gegenstände nicht nur beobachten oder damit Experimente anstellen, sondern darüber hinaus materiell tatsächlich in kleinste bestandteile zerlegen, oder auch umgekehrt aus solchen synthetisieren - die "klassischen" Naturwissenschaften wie Physik, Chemie, in vitro - Biologie.

Eine weitere Einteilungsachse wäre der Grad der Quantifizierung und Mathematisierung von Disziplinen. *Mathematisch-quantitative versus natur-sprachlich-qualitative Disziplinen.*

1.7 Zur Frage der Wertfreiheit der Wissenschaft

Bevor wir auf dieses 'heiß umstrittene' Thema eingehen, sei zunächst bemerkt, daß wir hier die Begriffe "Wert" und insbesondere "Norm" in engem Sinn verstehen. So verstehen wir unter "Normsatz" nur eine explizite Behauptungen, derzufolge ein Sachverhalt geboten bzw. gesollt ist, und zwar im ethischen Sinn. Eine rein *konventionelle* Definition oder Sprachregel wie "1 m soll die Länge des Urmeters in Paris sein" ist noch keine Norm in diesem Sinn. (Denn: daß das Maß "1m" gerade so und nicht so lang soll, ist überhaupt nicht ethisch wertvoll; ethisch wertvoll ist lediglich, daß man sich auf irgendein Maß einigt; und nur letzteres ist eine Norm in unserem eingeschränkten Sinn). Wir betonen dies, weil in gewissen Lagern der Philosophie der Begriff "normativ" ganz weit und unscharf verwendet wird, demzufolge dann alle nicht unmittelbar beobachtbaren oder empirisch reduzierbaren Theoriebestandteile und insbesondere sogenannte Definitionen gleich 'eine geheime normative Komponente haben' - eine Auffassung, die, wie wir später sehen werden (insb. Kap. 4), nicht nur eine Begriffsverwirrung, sondern auch ein echtes wissenschaftstheoretisches Erkenntnishindernis ist.

Zunächst ist zu bedenken, daß mit 'Wertfreiheit' der Wissenschaft unmöglich '*absolute* Wertfreiheit' gemeint sein kann, da die Wissenschaft ja selbst auf einen fundamentalen Wert anstrebt, das allgemeine wissenschaftliche Erkenntnisziel der Suche nach gehaltvollen Wahrheiten, bzw. damit zusammenhängende Unterziele. Wir sprechen hier von *wissenschaftsinternen* Werten. Man darf es nicht als paradox empfinden, daß die Wissenschaft zwar auf diesen internen Werten basiert, sie selbst aber natürlich auch nicht im fundamentalen Sinn begründen kann (daß gerade die stärksten theoretischen Systeme zumeist nicht alle ihre Voraussetzungen begründen können, ist eine durchaus gewöhnliche Erkenntnis moderner Wissenschaft). Die

wissenschaftsinternen Werte betreffen sozusagen die Entscheidung für die Wissenschaft; das wissenschaftlich Begründbare ist dann das, was aus dieser Entscheidung hervorgeht. - Was mit der Wertfreiheitsthese, nach dieser Betrachtung, jedenfalls nur gemeint sein kann, ist die Freiheit der Wissenschaft von *wissenschaftsexternen* Werten, das sind alle anderen Werte - natürlich tausendfach vielfältiger wie die wissenschaftsinternen - von persönlichen Motiven bis zu ökonomischen und politischen Interessen. Obwohl Max Weber seine Wertfreiheitsthese nicht immer so formulierte, geht es etlichen Stellen hervor, daß er sie in diesem auf externe Werte bezogenen Sinn gemeint hat.

Insofern Grundnormen und -werte nicht wissenschaftlich begründbar sind, so folgt schon natürlich bereits daraus, daß in wissenschaftlichen Begründungen der Rekurs auf Grundwerte bzw. die sie tragende Intuition nicht zulässig ist. Dennoch fließen an unzähligen Stellen des wissenschaftlichen Erkenntnisprozesses Wertungen ein, sodaß wir zeigen müssen, an welcher Stelle dieses Prozesses die Wertfreiheitsthese relevant wird, und in welcher Form. Gehen wir aus von der üblichen Einteilung des wissenschaftlichen Forschungsprozesses in drei Phasen, nämlich: *Entstehungszusammenhang*, (EZ) *Begründungszusammenhang* (BZ) und *Verwertungszusammenhang* (VZ). Ferner unterscheiden wir, wie erwähnt, zwischen den wissenschaftsinternen Werten (IW) und den wissenschaftsexternen Werten (EW), In (EZ) werden zuallerst auch die Problemstellungen definiert, der Bereich des Untersuchungsobjekts wird eingeschränkt, die Daten werden nach Relevanzgesichtspunkten selektiert. Dies ist sicherlich auch von wissenschaftsexternen Werten (EW) abhängig: welches Problem wichtig ist, entscheiden u.a. gesellschaftliche Interessen oder subjektive Interessen des Einzelforschers. Auch in (VZ) kommen selbstverständlich (EW) massiv ins Spiel - daß die Verwendung wissenschaftlicher Erkenntnisprodukte von gesellschaftlichen Interessen bestimmt wird, ist evident. Jedoch im Bereich (BZ) dürfen keine (EW) eine Rolle spielen - die Begründung wissenschaftlicher Sätze darf sich nur und ausschließlich an den Kriterien wissenschaftlicher Begründung orientieren, die wir in den Grundbestandteilen M1-M4 zusammenfaßten, und die sich auf das minimale Erkenntnismodell A1-A5 stützen, welches seinerseits als Mittel zur Realisierung des Wertes (IW) dient. Der Bereich (BZ) darf ausschließlich nur an den Werten in (IW) orientiert sein. Schematisch:

(EW), (IW)

(IW)

(EW),(IW)

(EZ)

(BZ)

(VZ)

Unsere "Wertfreiheitsthese" lautet also nicht

"in der Wissenschaft kommen keine Werte vor"

sondern lediglich

"in einem bestimmten Bereich der Wissenschaft, nämlich im Begründungszusammenhang, kommen keine wissenschaftsexternen Werte vor".

Dies ist eine viel schwächere, aber auch viel vernünftiger Fassung des Wertfreiheitspostulats als weiter obenstehende pauschale Form.

Eine wichtige Konsequenz obigen Schemas sei betont: die in (EZ) aufgrund (IW) und (EW) vorgenommene Selektion von untersuchten Parametern ist keinesfalls *erkenntnisneutral*, sondern wirkt indirekt auf den Begründungszusammenhang, insofern sie das, was als Resultat der wissenschaftlichen Analyse sich ergeben kann, einschränkt. Die in (EZ) vorgenommene Selektion muß also immer nur vorläufig sein und muß später zumeist korrigiert werden. So kann es sich im Verlauf der wissenschaftlichen Forschung als immer plausibler erweisen, daß zur Erfassung der relevanten Ursachen die Berücksichtigung von Parametern nötig sind, von denen in (EZ) bereits abstrahiert wurde (vgl. Kap. 3) Und das kann dann erkenntnishemmend sein, ja sogar zur 'wissenschaftlichen Ideologie' geraten: dann wäre tatsächlich auch (BZ) von (EW) 'irreversibel' beeinflusst, was - sofern von wirklicher wissenschaftlicher Begründung die Rede sein soll - natürlich nicht sein darf. Daraus folgt, daß die Selektion, die in (EZ) notwendigerweise geschehen muß - man muß sich ja einschränken und mit irgendetwas anfangen - selbst immer wissenschaftlich korrigierbar bleiben muß; woraus zugleich folgt, daß die Wissenschaft sich der externen Werte, die für ihre Problem- und Parameterselektion maßgeblich waren, auch im internen Begründungszusammenhang immer bewußt bleiben muß und diese transparent machen muß. Ein Beispiel: die klassische Schulmedizin war lange Zeit, z.T. noch heute, rein biologisch-physiologisch orientiert und abstrahiert von der menschlichen Psyche. Man kennt heute jedoch viele organische Krankheiten mit psychischen Ursachen, man erkennt die Bedeutung der psychosomatischen Interaktion. Ideologische Kämpfe zwischen Befürwortern und -gegnern entbrennen. Doch es handelt sich hier vorallem um eine theoretische Frage: und schon um sie überhaupt sachlich untersuchen zu können, muß man heute die Einschränkung auf biologische Parameter in medizinischen Modellen aufgeben und den Bereich des Psychischen miteinschließen.

1.8 Historischer Exkurs: Zur historischen Entwicklung der Wissenschaftstheorie

Der Beginn der modernen Wissenschaftstheorie setzt vorallem in den 30er Jahren des 20. Jahrhunderts mit der Philosophie des Wiener Kreises, dem sogenannten logischen Empirismus ein. Insbesondere seinen Nachfolger in den U.S.A., Rudolf Carnap und Carl Gustav Hempel, sowie der dem Wiener Kreis nahestehende Sir Karl Popper haben Grundlegendes geleistet. Ziel dieser Denker war es, durch Analyse der Naturwissenschaften zu einer wissenschaftlichen Philosophie zu gelangen. Die Werke von Carnap (1972, 1976), Hempel (1965, 1977) und Popper (1984) sind heute noch Standardwerke der Wissenschaftstheorie. Obzwar sich viele ihrer Thesen in ihrer ursprünglichen Form nicht mehr halten lassen, sind ihre Hauptgedanken und grundlegenden Problemstellungen auch heute noch maßgebend. (Natürlich wären auch viele andere Namen zu nennen, die für die Wissenschaftstheorie eine gewisse 'Vaterrolle' besitzen).

Dennoch hat es die Wissenschaftstheorie insofern immer schon gegeben, als immer schon Wissenschaftler und Philosophen über die Regeln der Wissenschaft nachgedacht haben. Wie die Philosophiegeschichte zeigt, waren die wissenschaftstheoretischen Auffassungen nie 'vom philosophischen Denkhimmel gefallen', sondern immer eng mit dem jeweiligen historischen Stand der Wissenschaftsentwicklung verbunden. Einige Beispiele: Eine ungeheure Faszination hatte in der Antike die Entdeckung des exakten mathematischen Denkens auf der Basis der geometrisch dargestellten Zahl (der Gründervater war Pythagoras, ca. 540 v. Chr.). Die Wissenschaftstheorie der sogenannten pythagoräischen Richtung vertrat demgemäß die Auffassung, das Wesen wissenschaftlicher Erkenntnis läge im Aufdecken exakter idealer mathematischer Zusammenhänge. Dadurch könnten die idealen und notwendigen Grundprinzipien alles Realen erfaßt werden - eine Auffassung, die insbesondere von Plato (427-348 v. Chr.) ausgeführt wurde. Empirische Beobachtung hatte dagegen für die wissenschaftliche Erkenntnis einen eher geringen Stellenwert, die wesentliche Erkenntnisgrundlage war Intuition, nicht empirische Beobachtung. Ziel der Wissenschaft war Erkennen der ewigen Ideen, und letztlich Eingehen der Seele in die Ewigkeit. Beides gilt auch noch für Aristoteles (384-322 v. Cr.), der zwar selbst extensiver Naturforscher war, dennoch glaubte, Naturgesetze würden im wesentlichen intuitiv erkannt werden. Beobachtungssätze und Falsifikation im modernen Sinn kommen in Aristoteles Logik (der Syllogistik) und in seiner Wissenschaftstheorie so gut wie nicht vor. Im Spätmittelalter gab es einen bedeutenden Aufschwung wissenschaftlichen Denkens. Bei Philosophen wie Roger Bacon (1210-1294), Wilhelm von Ockham (1295-1349) u.a. zeichnet sich bereits ein zunehmendes Mißtrauen in Möglichkeit der Erkenntnis von Naturgesetzen bzw. notwendigen Prinzipien auf der Basis intuitiver Spekulation ab. So gibt Robert

Grosseteste (1168-1253) eine sehr fortgeschrittene Darlegung der induktiven Methode - der Überprüfung von Gesetzhypothesen durch systematische Einzelbeobachtungen. Roger Bacon entwickelte eine Theorie des Experiments. Spätmittelalterliche Wissenschaftler, wie Jean Buridan (14. Jh.) und andere, haben Physik und Kosmologie weit über Aristoteles hinaustrieben und in gewissen Punkten sogar Galilei vorweggenommen. Insgesamt hatten die spätmittelalterlichen Wissenschaftler also bereits ein wesentlich ausgeprägteres Bewußtsein von der Fehlbarkeit der Erkenntnis als in der Antike. Dennoch war empirische Forschung - obwohl die spätmittelalterlichen Philosophen viel von der Naturbeobachtung sprachen - nicht ihr vordringliches wissenschaftliches Ziel; letzteres lag primär darin, zu zeigen, wie der Bereich der Naturerkenntnis mit dem Bereich der Religion zusammenhängt, und wodurch er sich unterscheidet.

Der entscheidende Entwicklungsschub der Wissenschaftstheorie vollzog sich erst in der Neuzeit, in der Auseinandersetzung mit den neuen Naturwissenschaften - deren Pioniere, wie Galilei (1564-1642), Johannes Keller (1571-1630), und vorallem (1642-1727), uns heute viel bekannter sind als ihre antiken und spätmittelalterlichen Vorfahren. Mit der neuzeitlichen Naturwissenschaft waren insbesondere drei Invasionen verbunden. (1.) Unter dem Eindruck ihres Erfolgs vertraten die Naturwissenschaftler sowie viele Philosophen mit Vehement jene charakteristische methodische Richtung, die man - 'pauschal' - den Empirismus nennt: alle wissenschaftlichen Hypothesen müssen durch Erfahrung überprüft werden; allein dies liefert Gründe ihrer Wahrheit. Rein, Geltung haben lediglich die Gesetze der Logik und Mathematik - doch rein Wahrheiten besagen nichts über die Realität, sondern beschreiben lediglich Strukturen exakten Denkens. (2.) Die Naturwissenschaft entwickelte eine empirische Überprüfungsmethode, die der bloßen empirischen Beobachtung weitgehend überlegen war, nämlich das wissenschaftliche Experiment. Darin werden nicht bloß jene Naturerscheinungen beobachtet, die die Natur von sich aus präsentiert, sondern die Natur wird vom Menschen durch die technische Herstellung von Experimentanordnungen gezielt in Situationen gebracht, in denen sie bestimmte und vorgegebene wissenschaftliche Fragen beantworten soll. Historisch war die Methode des Experiments eng verbunden mit der Entwicklung des Handwerks und dem Entstehen der Technik in der Renaissance. (3.) Zugleich vollzog sich eine intensive Weiterentwicklung logisch-mathematischer Methoden und präziser wissenschaftlicher Sprachen, die, im nachhinein betrachtet, oft erst den Durchbruch gewisser Problemlösungen ermöglichte. Beispielsweise wurde die Explikation des Begriffs einer kontinuierlichen Strecke wirklich befriedigend erst durch die Theorie der reellen Zahlen möglich; und die Begriffe der Momentangeschwindigkeit und -beschleunigung, mit denen man von der Antike bis ins Spätmittelalter hinein rang, wurden erst durch die Entwicklung des

Differentialkalküls durch Newton und Leibniz wirklich klar. - Und schließlich änderten sich in der neuzeitlichen Wissenschaft vorallem die Ziele der Wissenschaft: nicht länger war es ihr Ziel, die Seele des Menschen zum Himmel zu führen, so wie bei Plato, oder das Verhältnis des natürlichen Wissens zum Glauben zu klären, so wie im Spätmittelalter, sondern: möglichst viele empirisch gehaltvolle Wahrheiten zu sammeln. Damit verbunden war das strategische Ziel, die Natur in Form der Technik kontrollieren und zum maximalen Wohle des Menschen nutzen zu können.

Die Philosophie und Wissenschaftstheorie dieser Epoche teilte sich - um es historisch übervereinfacht zu sagen - in zwei Lager, das eine mehr im englischen, das andere mehr im deutschen Sprachraum beheimatet. Einerseits waren da die Empiristen, wie Francis Bacon (1561-1626), John Locke (1632-1704), David Hume (1711-1776), John Stuart Mill (1806-1873), um nur die wichtigsten zu nennen. Die gegenpolige philosophische Strömung soll durch zwei herausragende philosophische Richtungen angesprochen werden: der Rationalismus von René Descartes (1596-1650), und die Transzendentalphilosophie von Immanuel Kant (1724-1804). Descartes und Kant versuchten, in je verschiedener Weise, zu zeigen, daß auch in der empirischen Wissenschaft gewisse apriorische Prinzipien vorausgesetzt werden müssen, die über die reine Logik und Mathematik hinausgehen. Obwohl die philosophische Debatte darüber heute immer noch im Gange ist, muß doch gesagt werden, daß jedenfalls die Versuche von Descartes und Kant sowie viele weitere Versuche dieser Richtung, auf "synthetisch-apriorischen" Prinzipien empirische Hypothesen zu gründen, widerlegt wurden.

Insgesamt jedenfalls ist die moderne Wissenschaftstheorie mit dem weltgeschichtlichen Phänomen der empirischen Wissenschaften (wozu ja im 19. und 20. Jh. viele weitere, über die klassischen Naturwissenschaften hinausgehende Disziplinen kamen) aufs engste verbunden. Die zwei Hauptpfeiler, präzise wissenschaftlichen Sprache und Logik einerseits sowie und empirische Überprüfung andererseits, finden in der einen oder anderen Form, in den meisten Ausformungen gegenwärtiger Wissenschaftstheorie. Freilich haben aus moderner Sicht die meisten früheren Vertreter des Empirismus, inklusive des logischen Empirismus, eine viel zu enge und oft auch 'dogmatische' Konzeption des Empirismus vertreten. Z.B. dachte man anfänglich im logischen Empirismus, alle wissenschaftlichen Regeln müßten sich auf das Zusammenwirken von Beobachtungssätzen und Logik reduzieren lassen; alle wissenschaftlichen Begriffe müßten sich durch Definitionsketten auf Beobachtungsbegriffe zurückführen lassen. Derartige Aussagen wurden später bald widerlegt. Poppers Logik der Forschung war vergleichsweise liberaler: Theorien können ihm zufolge beliebig kreativ sein, sie können weit über das Beobachtbare hinausgehende Begriffsbildungen und Hypothesen enthalten. Wesentlich ist nur, daß

sie empirische Konsequenzen haben, anhand derer sie bewährbar und falsifizierbar sind (je nachdem, ob die Konsequenzen tatsächlich eintreten oder nicht). - Über diese tendenziellen Gemeinsamkeiten hinaus gibt es in der modernen Wissenschaftstheorie auch eine Reihe kontroversieller Fragen, von denen uns einige noch näher beschäftigen werden.

1.8.1 Verschiedene Positionen in Erkenntnis- und Wissenschaftstheorie - in Stichworten

Wissenschaftsauffassung in der Antike:

Pythagoras, Platon: Ziel = Erkenntnis des Wahren=Guten=Schönen

Es gibt ein Reich ewiger Ideen, Ideale, von denen die empirische Welt ein unvollkommenes Abbild ist.

Vorbild sind mathematische Ideen, z.B. Proportionen, die sich in der Geometrie wie in der Musik finden (auch noch Kepler).

Am mathematischen Modell orientiert sich Ethik und Ästhetik; umgekehrt dient Mathematik dem Schönen.

Erkenntnis: wesentlich durch Intuition; Ablehnung bloßer Beobachtung um ihrer selbst Willen; daran sollen letztlich Tugenden erkannt werden.

Aristoteles: keine Annahme eines Reiches ewiger Ideen.

Wahrheit = Korrespondenztheorie. induktiv-deduktive Methode. Logik.

Aber auch bei ihm: Schönheit geht ein in Wahrheit: vollkommene Bahnen der Gestirne, "weil Gott sie schuf".

Viel Naturbeobachtung, aber naive Beispielmethode: einige -> alle.

Höchste Form der Induktion ist intuitive Wesensschau, Evidenzerlebnis.

Kein Fallibilismus. Kaum Erkenntniskritik.

Nach Aristoteles, Wissenschaft in Alexandrien: Ptolemäus, Archimedes; vorher schon: Euklid, u.a.m.

Trennung Wissenschaft von Ethik und Metaphysik, Theologie.

Empirische Phänomene -- Erklärung der Empirie. Philosophie wird mystisch-theologisch.

Pythagoräer versus Empiristen

Mittelalter bis Hochscholastik: Wissenschaft im Dienst der Theologie.

Erkenntnis des Wahren = Erkenntnis Gottes, und damit zugleich des Guten und Schönen.

(Hoch bis) Spätscholastik (13., 14. Jh.) : Vorwegnahme der naturwissenschaftlichen Methode:

--> ausgelöst durch Bekanntwerden des aristotelischen Schrifttums, das verunsicherte weil es von Kirchenlehre bis dato abwich

--> vornehmlich ausgelöst durch die Philosophen des Franziskanerordens, s.u., welche daraus nicht dogmatische Konsequenzen (Ketzerium und Inquisition) sondern kritische Konsequenzen zogen.

Robert Grosseteste: Methode der Falsifizierung

Roger Bacon: empirische-experimentelle Methode, technische Visionen

John Duns Scotus: induktive Methode (Übereinstimmung und Unterschied)

William v. Ockham: Fallibilismus; Ockhams Rasiermesser

Jean Buridan: Impetustheorie, Vorwegnahme Galileis; Überwindung der Paradoxien der Aristotelischen Physik.

In der Renaissance: Romantische Gegenteilstendenz gegen MA-scholastische Theorie. Aber Übernahme ihrer Lehren; Übersetzung ins Deutsche, Verbreitung durch Buchdruck --- Verbindung Theorie mit den Handwerkern-Künstlern. Erst hier entsteht *die experimentelle Methode, die mehr als bloßes Spiel zu Ehren Gottes ist.*

Pioniere der Naturwissenschaft: Galileo Galilei, Johannes Kepler, Isaac Newton

Parallel dazu entwickelt sich die *Philosophie des Empirismus:* (16.-17.Jh.)

Francis Bacon: Wissenschaftspropagandist, Betonung des Experiments gegenüber bloßer Beobachtung. kein Fallibilist, Dogmat. "Machbarkeitsphilosophie".

Ziel ist nun: Wahrheit + *Kontrolle*

John Locke (17. Jh.): Realist, primäre versus sekundäre Qualitäten; naiver Empirismus und Psychologismus der Ideenassoziationen. Logik fehlt. Notwendiges Wissen (kein Fallibilist).

David Hume (18. Jh.): Kritik an Beweiskraft der induktiven Methode (Gewohnheit) und an naivem Kausalitätskritik. Fallibilismus.

John Stuart Mill (19.Jh.): Fortschritte in der Logik. Präzisierung der induktiven Methode (Übereinstimmung und Unterschied). Glaubt an ihre Beweiskraft.

Gegenposition in Philosophie: *Rationalismus:*

Rene Descartes (17. Jh.): Trotz methodischer Skepsis "cogito ergo sum" glaubt er an apriori-Wissen über die Realität durch Intuition und Evidenz. Apriori-Physik: widerlegt.

Gottfried W. Leibniz (17. Jh.), glaubt ebenso an notwendiges Wissen und angeborene Ideen, die notwendigerweise wahr sind.

Immanuel Kant (18. Jh.), versucht Synthese von Rationalismus und Empirismus in Form der *Transzendentalphilosophie*. Lenkt Augenmerk auf *Grenzen der Erkenntnis*. Es gib bei ihm Synthetica apriori.

Strömungen der Gegenwart 19. Jh. und vornehmlich 20. Jh.:

Die moderne logisch-empirische Wissenschaftstheorie:

Vorläufer: *Begründer der modernen Logik:* Boole, Peirce, Frege, Russell, Tarski (19/20. Jh.)

Naturwissenschaftler: Pierre Duhem (19/20.Jh.) --- Holismus der Falsifikation, Struktur wissenschaftlicher Theorien

Ernst Mach, Henri Poincare (Konventionalismus), Percy Bridgman (Operationalismus)

Wiener Kreis: Moritz Schlick, Otto Neurath, Viktor Kraft, Rudolf Carnap, Carl Gustav Hempel alle anfangs zu sehr empiristisch; siehe die Diskussion um das Sinnkriterium: Verifizierbarkeitskriterium, dann Definierbarkeitskriterium...

vgl. den *Behaviorismus* in der Psychologie

- im Gegensatz zu Karl Popper

Berliner Kreis: Hans Reichenbach, sowie von Mises: *statistische Methode* moderne *Philosophy of Science*, zunächst im angelsächsischen Sprachraum
In der Soziologie: Max Weber "Wissenschaft als Beruf".

Der *Pragmatismus:* Charles S. Peirce: ganz ähnlich, bloß:

-- Konsensustheorie der Wahrheit;

-- sucht auch nach *Regeln der Entdeckung* (hat er von William Whewell)

Der Einbruch des *Relativismus* durch die Wissenschaftsgeschichte:

Norwood R. Hanson: "Interpretationen" von Beobachtung

Thomas S. Kuhn: ebenso

Paul Feyerabend: Einbruch der Wertungen, des Irrationalismus

Theorieabhängigkeit der Beobachtung --> Subjektivität, weil gefangen im theoretischen Paradigma

Kultur und Werte fließen in Beobachtung ein --> gefangen in Kultur und Weltanschauung

Zwei Arten des Relativismus: (= Anti-Objektivismus)

1. Es gibt eine Realität, aber keine objektive Erkenntnis davon (Epistemischer Relativismus)

2. Es gibt "soviele Realitäten wie Weltanschauungen" (Ontologischer Relativismus)-
-> Carlos Castaneda, Hans Peter Dürr, auch Feyerabend: "der Regentänzer und die Wahrheit...".

Moderner *Rationalismus* oder *Transzendentalphilosophie:* Erkenntnis benötigt

apriori-Annahmen.- Z.B. Karl-Otto Apel (Peirce-Rekonstruktion).

"Kritischer Rationalismus" von Karl Popper ist eigentlich kein Rationalismus
Moderne Metaphysik: *Saul Kripke*, ...

Konstruktivismus: Hat in der Philosophie verschiedentliche Rolle gespielt. "Wirklichkeit sei konstruiert". Siehe Wahrnehmungstäuschungen. Vertreter äußern sich oft sehr unklar über epistemologischen versus ontologischen (=radikalen) Konstruktivismus. Dabei:

Gemäßigte (epistemische) Variante: Unsere Sicht der Wirklichkeit ist konstruiert. Es gibt aber "Dinge da draußen", eine unabhängig existierende Realität. (Klaus Holzkamp, Jean Piaget)

Radikale (ontologische) Variante: Auch die Realität selbst ist konstruiert (Ernst von Glasersfeld).

Variante dazwischen: Was die Realität "an sich" ist, kann man nicht sagen (Piaget, Glasersfeld; streng: die philosophischen Konstruktivisten: Hugo Dingler, Paul Lorenzen).

Geisteswissenschaften, Hermeneutik:

Wilhelm Dilthey, Wilhelm Windelband, (19/20. Jh.), Hans Gadamer:

-- Gegenstand der Gw = menschliche Handlung, menschliche Sprache
kein physikalischer Gegenstand, sondern Interpretation, Bedeutung, Sinn.
Harmlose Variante: Intersubjektivität, kein Widerspruch zu moderner
Auffassung!).

Radikalere Variante: Verstehen prinzipiell subjektiv --- Relativismus.

Z.B.: Alfred Lorenzer: In Psychoanalyse

Z.RT. auch "qualitative Forschung"

-- Man sucht nicht nach *Gesetzen*, sondern Beschreibung des Einzelfalls.
(Auch das läßt sich in induktiver Einzelfallerklärung nachvollziehen).

-- Man soll eine weiche metaphorische Sprache benutzen: gegen *Logik iwS*
"qualitative Forschung" ---- empirische Überprüfung wird fraglich.

-- Methode ist *Einfühlung, Verstehen durch Emphatie*: Ist wissenschaftlich
unkritische Vorgehensweise, weil nicht nachvollziehbar, nicht reproduzierbar.

Quantitativ versus qualitativ: Ein *Sammelsurium* verschiedener Kritiken und
Auffassungen, Relativismus, kritische Theorie (s.u.), Hermeneutik, denen bloß eines
gemeinsam ist: Kritik an der empirisch-analytischen Auffassung.

Kritische Theorie: Theodor W. Adorno, (Neomarxismus), Jürgen Habermas (der
frühere): *Erkenntnis allemal interessegebunden* -- hier kein "Mangel" sondern
positiv zu begreifen: Erkenntnis muß zugleich politisch sein. (Vgl. auch Marxismus).

Hier: Wahrheit = Erkenntnis des Guten zugleich,

Humanistische Psychologie: Abraham Maslow: Es geht in Forschung zugleich um Wert und Würde des Menschen, um Erfassung des Sinns des Lebens -- auch hier: Wahrheit + das Gute.

Erläuterungen zur folgenden Übersicht:

Es wird hier für die wichtigsten Positionen angeführt, welche unserer Annahmen A1-A5 und methodischer Merkmale M1-M4 zutreffen. Zusätzlich nehmen wir die Wertneutralitätsthese als zusätzlichen Parameter WN hinzu.

Beim Realismus wird zwischen naivem (n) - und minimalem (+) unterschieden; ebenso beim Empirismus zwischen naivem (n+) und minimalem (+).

"?" bedeutet, daß es hier in dieser Position ein "für und wider" gibt. ?+ bedeutet eher plus; ?- bedeutet eher minus. *Natürlich ist die Tafel äußerst ungenau*, was die Details betrifft, sie gibt nur eine Groborientierung, die das Detailstudium der Positionen nicht ersetzt!!

Festgehalten wird in dieser Tafel, was die Positionen selbst vertreten -- ich denke, daß viele davon eigentlich inkonsistent sind und *implizit* was anderes voraussetzen.

Position	WN	A1	A2	A3	A4	A5	M1	M2	M3	M4
Pythagoras/Plato	-	n+	-	+	?	-	+	-	-	-
Aristoteles	-	n+	-	+	+	-?	+	-?	+	?+
Alexandrien	+	n+	-	+	+	n+	+	+	+	?+
MA bis 12 Jh.	-	n+	-	+	?	-	+	-	+	-
Spätscholastik	?+	+	?+	+	?+	n+	+	+	+	?+
<u>Empirismus:</u>										
Bacon, Locke	?+	n+	-	+	-	n+	+	+	+	?+
Hume	+	-	+	+	-	+	+	+	+	+
Mill	?+	+	-	+	+	n+	+	+	+	+
<u>Rationalismus</u>	?-	+	-	+	+	-	+	?-	+	?-
<u>Kant</u>	-	+	?-	+	+	?	+	+	+	+?
<u>Anal.-emp.Wth:</u>										
frühe	+	?	+	+	+	n+	+	+	+	+
postpositivistische	+	+	+	+	+	+	+	+	+	+
<u>Pragmatismus</u>	+	?-	+	+	+	+	+	+	+	+
<u>Relativismus</u>										
gemäßigt	?	?+	+	-	-	-	+	?	+	-
radikal	?	-	+	-	-	-	+	?	+	-
<u>Moderne Transz.</u>	?	+	-	+	-	?	+	?	+	?
<u>Konstruktivismus</u>										
gemäßigt	+	+	-?	+	+	+	+	+	+	+
radikal	?	-	?	-	?	-	+	-	+	-
<u>Gw-Hermeneutik</u>										
gemäßigt	?-	?+	+	+	-	+	-	+	?	+
radikal	-	-	?	-	-	-	-	?+	-	-
<u>Kritische Theorie</u>	-	-	+	-	-	-	?+	?+	+	?+

1.9 Literaturhinweise

Zu Kap. 1.2: Eine wissenschaftshistorische Einführung in die Wissenschaftstheorie: Losee, J. (1977), *Wissenschaftstheorie. Eine historische Einführung*, C. H. Beck, München. Eine gute Wissenschaftsgeschichte ist: Crombie, A. (1977): *Von Augustinus bis Galilei*, dtv, München (Orig. 1959). Zum Wiener Kreis: Schleichert, H. (Hg., 1975): *Logischer Empirismus - der Wiener Kreis*, W. Fink, München.

Zu Kap. 1.3-1.4 Klassische wissenschaftstheoretische Werke von Carnap, Hempel und Popper sind: Carnap, R. (1976), *Einführung in die Philosophie der Naturwissenschaft.*, Nymphenburger Verlagsbuchhandlung, München (3. Auflage). Hempel, C. G. (1977): *Philosophie der Naturwissenschaften*, dtv, München (3. Aufl.). Popper, K. (1984), *Logik der Forschung*, Mohr, Stuttgart (8. Aufl.).

Modernere wissenschaftstheoretische Einführungswerke: Bunge, M. (1967): *Scientific Research*, 2 Bde., Springer, Berlin. Kutschera, F. v. (1972), *Wissenschaftstheorie , Bd. I und II*, W. Fink (UTB), München. Stegmüller, W. (1969-1986), *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*, 4 Bände, Springer, Berlin-Heidelberg-New York. Weingartner, P. (1978): *Wissenschaftstheorie, Band 1*, frommann-holzboog, Stuttgart (2. Aufl.). Vgl. auch Schurz, G. (1983): *Wissenschaftliche Erklärung*, dbv, Graz, Kap. I und II.

Die zitierten Werke von Kuhn und Feyerabend: Kuhn, T. (1976): *Die Struktur wissenschaftlicher Revolutionen*, 2. rev. Aufl., Suhrkamp, F/M. Feyerabend, P. (1976): *Wider den Methodenzwang*, Suhrkamp, F./M.

Zur Rechtswissenschaft: Savigny, Eike v. (1976): *Juristische Dogmatik und Wissenschaftstheorie*, Beck, München.

Zum Werturteilsstreit: Albert, H./Topitsch, E. (Hg, 1971): *Werturteilsstreit*, Wiss. Buchgesellschaft, Darmstadt. Vgl. Max Webers klassisches Werk: *Gesammelte Aufsätze zur Wissenschaftslehre* , Mohr, Tübingen, 1968, 3. Aufl.

2. Die logischen Grundlagen

2.1 Einteilung von Begriffsarten

Begriffe lassen sich in mehreren Dimensionen klassifizieren. Erstens nach ihrem "logischen Typ", also ob es sich um singuläre Terme oder generelle Begriffe oder gar rein logische Begriffe handelt. Zweitens nach ihrem "Inhaltstyp", zum Beispiel ob es sich um deskriptive oder um Wertbegriffe handelt. Drittens nach ihrer "Abgestuftheit" bzw. ihrer "Metrik", also ob es sich z.B. um qualitative oder um quantitative Begriffe handelt.

2.1.1 Einteilung nach dem logischen Typ:

Singuläre Terme (Nichtlogische Begriffe)	(bezeichnen einzelnen Individuen oder individuelle Situationen)
Generelle Begriffe	
Prädikate	einstellige - bezeichnen Eigenschaften Merkmale, oder Arten
	mehrstellige - bezeichnen Relationen
Funktionsbegriffe	- bezeichnen Funktionen
Logische Begriffe	wahrheitsfunktionale Satzverknüpfungen, nicht, und, oder, wenn-dann; Quantoren, wie "für alle", "es gibt"; intensionale Satzoperatoren wie "notwendig", "möglich", "wahrscheinlich" Prädikatoperatoren wie die statistische Wahrscheinlichkeit
	Variablen (für Individuen, Prädikate, usw.)

Die singulären Terme und generellen Begriffe nennen wir *Nichtlogische Begriffe*.

Erläuterungen:

Nichtlogische Begriffe -- und die in der Logik verwendeten Zeichen dafür:

1. *Individuenkonstanten*: a,b,c,...; auch indiziert a_1, a_2, a_3, \dots Sie bezeichnen Einzeldinge, also Individuen verschiedenster Art. In der natürlichen Sprache

entsprechen ihnen Eigennamen, also z.B. "Peter", aber auch für persönliche Fürwörter wie "ich", "er", oder hinweisenden Fürwörter wie "das Ding hier und jetzt", etc. Spezielle Individuenkonstanten sind Ortskonstanten $s_1...$ und Zeitkonstanten $t_1...$

2. *Prädikate*: F, G, \dots ; F_1, F_2, \dots . Sie bezeichnen allgemeine Merkmale, die auf viele verschiedene Individuen zutreffen können. Einstelligen Prädikate werden nur auf eine Individuenkonstante angewandt und bezeichnen Eigenschaften oder Arten. Mehrstellige Prädikate heißen auch Relationsprädikate und bezeichnen Beziehungen (Relationen) zwischen mehreren Individuen. Die Anwendung eines n-stelligen Prädikates auf n Individuenkonstanten (oder -variablen) ergibt einen atomaren Satz. Einige Beispiele von Formalisierungen mittels atomarer Sätze (der Strich "-" zeigt dabei die Übersetzungsbeziehung zwischen formalem Symbol und natürlichsprachigem Begriff bzw. Satz auf):

F - die Eigenschaft, traurig zu sein; zumeist schreibt man dies so: Fx - x ist traurig
 a - Peter

Ergibt folgende Formalisierung: Fa - Peter ist traurig.

Rxy - x übt auf y eine Kraft aus

a - diese Kompaßnadel

b - der Nordpol

Dies ergibt: Rba - der Nordpol übt auf diese Kompaßnadel eine Kraft aus.

3. *Funktionszeichen*: f, g, \dots ; auch f_1, f_2, \dots . Funktionszeichen können, wie Prädikate, einstellig oder mehrstellig sein. Sie ordnen ein- oder mehreren Individuen eindeutig ein weiteres Individuum bzw. eine individuelle Größe zu. Z.B. mit f - Mutter von; a - Peter bezeichnet " fa " die Mutter von Peter. f ergibt also in Anwendung auf a jene Person, die die Mutter von a ist. Funktionszeichen kennt man aus der Mathematik: z.B. ordnet das zweistellige Funktionszeichen "+" je zwei Zahlen ihre Summe zu: " $x+y$ ". Funktionszeichen sind insbesondere zur Formalisierung von quantitativen wissenschaftlichen Begriffen wichtig. Z.B. ist "Masse" eine Funktion von Körpern; d.h. $m(a)$ bezeichnet die Masse des Körpers a - d.h. eine bestimmte reelle Zahl mit einer bestimmten physikalischen Einheit.

Logische Symbole:

4. *Aussagenlogische wahrheitsfunktionale Satzoperatoren*: \neg für die Negation (nicht); \wedge für die Konjunktion (und); \vee für die Disjunktion (das einschließende oder),

→ für die materiale Implikation (das materiale wenn-dann), ↔ für die Äquivalenz (das materiale genau wenn, dann). D.h. für "¬A" lies "nicht A", für "A∧B" - "A und B", für "A∨B" - "entweder A oder B oder beides", für "A→B" - "wenn A dann B" (im Sinne von: "falls nicht A, dann egal ob B; falls aber A dann gewiß B"), usw.

5. *Individuenvariablen:* x, y, z ...; auch indiziert x_1, x_2, x_3, \dots bezeichnen variable Einzeldinge, und werden insbesondere im Zusammenhang mit Quantoren wichtig (s. 7).

6. *Quantoren:* 6.1 \forall - der Allquantor: " $\forall x$ " bedeutet "für alle x (eines zugrundegelegten Objektbereichs, meist der universale Bereich aller Dinge) gilt:..." 6.2 \exists - der Existenzquantor; " $\exists x$ " bedeutet "es gibt mindestens ein x (im zugrundegelegten Objektbereich), für das gilt:...". Indem man an die Ausdrücke " $\forall x$ " bzw. " $\exists x$ " prädikatenlogische Sätze mit gebundenen der Individuenvariable x anfügt, enthält man die elementarsten Typen quantifizierter Sätze der Prädikatenlogik. Beispiele:

$\forall x(Fx \rightarrow Gx)$ Für alle x gilt: wenn x F ist, ist x G; oder kurz: Alle F sind G. Z.B. Fx - ein Metall zu sein, Gx - Strom zu leiten. Das ergibt: $\forall x(Fx \rightarrow Gx)$ - alle Metalle leiten Strom.

$\exists x(Fx \wedge Gx)$ - Es gibt ein x, für das gilt: x ist F und G; kurz: es gibt ein F, das G ist. Z.B. Fx - x ist ein Mensch, Gx - x ist älter als 150 Jahre. Das ergibt: $\exists x(Fx \wedge Gx)$ - Es gibt einen Menschen, der älter als 150 Jahre wurde. Bzw. $\neg \exists x(Fx \wedge Gx)$ - es gibt keinen (=nicht: es gibt einen) Menschen, der älter als 150 Jahre wurde.

Quantoren können auch hintereinander (verschachtelt) auftreten:

$\forall x(Gx \rightarrow \exists y(Ty \wedge Sxy))$ - Für alle x, die die Eigenschaft G haben, gibt es ein y, das T ist, und zu x in Beziehung S steht. Mit Gx - x ist ein Gas, Tx - x ist eine gewisse Temperatur und Sxy - x schmilzt zu y besagt dies z.B.: Alle Gase besitzen eine Temperatur, zu der sie schmelzen; kurz: Alle Gase haben einen Schmelzpunkt.

Mit $f(x,t)$ - die auf Körper x zum Zeitpunkt t ausgeübte Gesamtkraft, $m(x)$ - die Masse von x, $a(x,t)$ - die Beschleunigung von x zu t, und \cdot - die mathematische Multiplikation lautet das 2. Newtonsche Gesetz formalisiert so: $\forall x \forall t(f(x,t) = m(x) \cdot a(x,t))$ - für alle x und alle t: die auf x zu t ausgeübte Gesamtkraft ist gleich dem Produkt der Masse von x und der Beschleunigung von x zu t.

7. *Intensionale Satzoperatoren:* sie besitzen als Argumentstellen ein oder mehrere Sätze, und zusätzlich evtl. Individuenkonstanten. Die wichtigsten Typen:

(7a) Modale Operatoren ... - einstellig, ein Satz als Argument. N - es ist notwendig, daß. und M - es ist möglich daß. Also z.B.: $M(Fa)$ - es ist möglich, daß Peter traurig ist.

(7b) Deontische Operatoren (normative Operatoren) und Wertoperatoren - einseitig, ein Satz als Argumentstelle. G - es ist geboten, daß (gesollt, daß); E - es ist erlaubt, daß; WO - es ist wertvoll, daß.

u.a.m.

8. *Wahrscheinlichkeit*

8a. Der statistische Wahrscheinlichkeitsfunktork: Er wird auf ein Prädikat bzw. eine Klasse F angewandt und liefert in Anwendung auf die Klasse F die relative Häufigkeit der F-Individuen. $p(F)=0,8$ (bzw. =80%) besagt also, daß 80% aller Individuen (eines festgelegten Objektbereichs, der hier auch Grundgesamtheit heißt) die Eigenschaft F haben. $p(G/F)$ bezeichnet die sogenannte bedingte Wahrscheinlichkeit, d.h. die relative Häufigkeit der G-Individuen innerhalb der Klasse der F-Individuen. $p(G/F)=0,7$ besagt also, kurz formuliert, daß 70% aller F's auch G's sind.

8.b Eine andere Art von Wahrscheinlichkeitsfunktork ist die epistemische (oder subjektive) Wahrscheinlichkeit w , die auf Sätze angewandt wird und deren epistemischen Glaubensgrad bezeichnet. Wir werden hier davon kaum Gebrauch machen.

Zur Bedeutung des logischen Formalisierens für die Wissenschaftstheorie

Die Formalisierung eines wissenschaftlichen Satzes besteht darin, ihn ganz oder teilweise in eine formale bzw. *zumeist halbformale* Modellsprache zu übersetzen. Dies ist in der Wissenschaftstheorie aus mehreren Gründen wichtig. Ein erster Zweck logischen Formalisierens ist es, den logischen Typ von Begriffen und (voralledem) Sätzen zu erkennen, z.B.: handelt es sich um einen Existenzsatz oder einen Allsatz? - davon hängt nämlich zu einem Gutteil die wissenschaftstheoretische Natur des Satzes ab. Ein zweiter und wohl der Hauptzweck logischen Formalisierens ist, die logische Gültigkeit von natursprachlich formulierten Argumenten zu überprüfen. Denn ist einmal die Formalisierung gelungen, so kann die Frage, ob ein Argument gültig ist, nach den elementaren Regeln logischen Schließens entschieden werden. Dies ist z.B. wichtig, um zu überprüfen, ob die Behauptung, eine Theorie liefere eine deduktive oder auch nur statistische Erklärung eines Sachverhaltes, tatsächlich zutrifft. Meistens entdeckt man dabei, daß in der natürlichsprachigen Formulierung gewisse stillschweigende Voraussetzungen gemacht wurden, die für die Gültigkeit des Arguments jedoch unerläßlich sind - und eben darin liegt der dritte Zweck: logisches Formalisieren ist daher eine wichtige Hilfe, um unhinterfragte Voraussetzungen zu erkennen. Einen vierten Zweck schließlich, nämlich die Herausarbeitung einer interdisziplinär übergreifenden präzisen Sprache, haben wir bereits erwähnt.

Um logische Zusammenhänge zu erkennen ist nicht jedesmal die Übersetzung

natürlichsprachiger Sätze in formale nötig. Doch zum einen ist es für den Anfänger eine nützliche *Übung*, um schnell logische Beziehungen zu erkennen, und zum zweiten ist es gerade in philosophischen und wissenschaftstheoretischen Kontexten - wo es sich z.T. um sehr abstrakte Fragen handelt, die leicht in 'sprachliche Fallen' bzw. Fehlschlüsse führen - sehr wichtig.

2.1.2 Einteilung von Begriffen nach dem Inhaltstyp

Diese Einteilung schließt an die vorige an und betrifft die Einteilung der *nicht-logischen Begriffe*.

(logische Begriffe)

deskriptive Begriffe

Beobachtungsbegriffe

theoretische Begriffe

präskriptive (ethische)
Begriffe

Normbegriffe

Wertbegriffe

Erläuterungen:

Über die Unterscheidung deskriptive versus präskriptive oder allgemeiner ethische Begriffe sprachen wir bereits.

Die Unterscheidung zwischen Beobachtungs- und theoretischen Begriffen ist wissenschaftstheoretisch eminent wichtig. Theoretische Begriffe sind solche, die nichtwahrnehmbare Merkmale bezeichnen, nicht mit den 5 Sinnesorganen beobachtbare Merkmale - natürlich können theoretische Merkmale mit Meßinstrumenten meßbar sein. Theoretische Begriffe werden aus der Beobachtung "indirekt" erschlossen. Natürlich ist die Unterscheidung auch problematisch und unscharf. Wir gehen darauf in Kap. 4 ein. beispielsweise sind "Rot", "Eckig", in der Physik "Ort, Zeit, Geschwindigkeit" Beobachtungsbegriffe; dagegen sind "Kraft", "elektrisches Feld", oder in der Psychologie "Charakter", "Motivation" etc. theoretische Begriffe.

Unter Normbegriffen verstehen Begriffe wie "es ist geboten, daß", bzw. "verboten", "erlaubt", "gesollt", usw. Unter Wertbegriffen z.B. "es ist (ethisch) wertvoll, daß", "es

ist gut, daß", usw. Wertbegriffe sind oft komparativer Natur, oft auch quantitativer Natur, z.B. Nützlichkeitswerte. (Zur Metrik s. u.).

1.2.3 Einteilung der Begriffe nach ihrer Metrik bzw. nach ihrem Skalentyp

qualitative Begriffe	Klassifikatorische Begriffe Nominalskala
	Komparative Begriffe Ordinalskala
quantitative Begriffe	Intensive Größen Intervallskala
	Extensive Größen Verhältnisskala

Erläuterungen:

1) *Qualitative Begriffen bzw. Nominalskalen (Kategorialskalen)*: Hier klassifiziert man Objekte nach verschiedenen Eigenschaften bzw. Merkmalen, die sich gegenseitig ausschließen und zusammengenommen den Möglichkeitsspielraum erschöpfen (man sagt, die Klassifikation ist disjunkt und exhaustiv); mehr Beziehungen bestehen nicht. Der einfachste Fall ist die zweiwertige Begriffsfamilie (dichotomische Variable), z..B. x ist ein Mensch oder nicht ein Mensch; x ist in psychiatrischer Behandlung oder nicht; usw. Die meisten qualitativen Begriffe der natürlichen Sprache sind von dieser Art. Es gibt aber auch mehrwertige Nominalskalen bzw. Begriffsfamilien. Z.B.: die Farbbegriffe: rot, grün...; die Wahlmöglichkeiten bei politischen Wahlen: x hat Partei x, y, z, oder ungültig oder nicht gewählt. Soziale Schichten: Unterschicht, Mittelschicht und Oberschicht. - Die Zuordnung von Nominalskalen zu Zahlen ist im Prinzip überflüssig, da nur qualitativ unterschieden wird.

2) *Komparative Begriffe bzw. Rangskalen (Ordinalskalen)*. Hier wird zusätzlich erfaßt, ob ein Merkmal stärker oder weniger stark ausgeprägt ist. Z.B. können wir beim subjektiven Hörempfinden beurteilen, welcher Ton lauter empfunden wurde als der andere, ohne genau sagen zu können um wieviel der Ton lauter war. Am mathematisch bequemsten formalisiert man ein rangskaliertes Merkmal als eine Funktion von x in die Menge natürlicher Zahlen; letztere geben dann die Rangplätze an. Wesentlich bei Rangskalen ist aber, daß die Objekte hinsichtlich der

Merkmalsausprägung zwar geordnet werden können, aber die 'Distanzen' zwischen den Rängen unbekannt sind. Das Weber-Fechnersche Gesetz besagt beispielsweise, daß zwischen physikalischer Lautstärke und subjektiver Hörempfindung kein linearer, sondern ein logarithmischer Zusammenhang besteht; würde man unsere subjektive Rangeinteilung hier als objektives Distanzmaß interpretieren, so hätte man einen Fehler begangen. Typische Rangskalen in den Sozialwissenschaften sind Einstellungsmessungen von Personen, z.B. in Interviews: finden Sie das Umweltproblem gefährlich?: 1,.....,6 (überhaupt nicht, ..., sehr), usw. Wichtige Methodenfragen der Sozialwissenschaften in diesem Zusammenhang, auf die wir hier nicht näher eingehen, betreffen z.B. die Tatsache, daß faktische Einstellungen oft nicht transitiv sind, so wie es die Axiome einer Rangskala fordern, oder die Frage, unter welchen Bedingungen sich eine Rangskala statistisch wie eine Intervallskala behandeln läßt.

3) *Quantitative Begriffe; Intervallskalen (Differenzskalen) und Verhältnisskalen (Ratioskalen)*. Bei der Messung psychischer Merkmale kommt man nur schwer über Rangskalen hinaus, obzwar es hier verschiedene Kriterien gibt, wann man die Skala 'wie eine Intervallskala' behandeln kann. Echt quantitative Begriffe, wo man den Ausprägungsgrad direkt, z.B. durch einen Zeiger messen kann, findet man typischerweise in den Naturwissenschaften. Formal sind quantitative Begriffe Funktionen von Objekten in die Menge der reellen Zahlen. Man unterscheidet zwischen Intervallskalen, wo man nur die Intervalle bzw. Differenzen zwischen einzelnen quantitativen Ausprägungen objektiv messen kann, aber keinen Nullpunkt der Skala kennt, und Verhältnisskalen, wo man zusätzlich einen Nullpunkt kennt. Intervallskalen sind z.B. Ort und Zeit - objektiv kann man hier nur Distanzen messen; es gibt weder einen absoluten Nullpunkt des Ortes oder der Zeit. Man kann also nicht sagen "das Jahr 2000 ist doppelt so spät wie das Jahr 1000" - das ist relativ auf Christus Geburt; wurde man etwa vom Steinzeitmenschen aus messen; so hätten man z.B. 50100 und 502000. Verhältnisskalen sind physikalische Größen wie Masse, Volumen oder Kraft - hier gibt es einen absoluten Nullpunkt; hier macht es daher auch Sinn, zu sagen, der 100kg schwere Peter wiegt doppelt so viel wie sein 50 kg schweres Weib, etc. - Quantitative Begriffe findet man auch in der Ökonomie - auch Bevölkerungsanzahlen, Geld- und Gütermengen sind Verhältnisskalen.

Unter dem *Metrisierungsproblem* (auch Meßproblem) versteht man allgemein die Frage, ob die mathematische Struktur der messenden Skala (numerisches Relativ) der faktischen Struktur des quantitativen empirischen Merkmalsausprägung (empirisches Relativ) angemessen ist. Die Beantwortung dieser Frage ist keinesfalls bloß 'konventionell', sondern von theoretischen Annahmen über zu messenden Objektbereich abhängig (z.B. von Annahmen über die 'Rationalität' von Personen bei Einstellungsmessungen unter Voraussetzung der Transitivität: a größer b und b

größer c \rightarrow a größer c). Durch inadäquate Meßskalen, deren Numerik feiner skaliert als das gemessene empirische Merkmal strukturiert ist, kann man Artefakte erzeugen (wie obiges Beispiel "2000 n.C. ist doppelt so spät wie 1000 n.C.").

2.2 Klassifikation von Satzarten

Die Klassifikation der Sätze setzt die Klassifikation der Begriffe voraus und nimmt auf sie Bezug. Auch Sätze werden wir nach mehreren Dimensionen klassifizieren - nach ihrem Inhaltstyp, nach ihrem Allgemeingrad, und nach ihrem methodologischen Status.

2.2.1 Klassifikation von Satzarten nach dem Inhaltstyp

	logisch determiniert	
analytisch	definitiv determiniert	
		empirisch
	deskriptiv	theoretisch
synthetisch	präskriptiv	
		Normsätze
		Wertsätze
	(gemischt)	

Erläuterungen:

1. *Analytische Sätze* sind solche, deren Wahrheitswert unabhängig von der empirischen Beschaffenheit der Welt allein durch die Regeln des logischen Schließens oder durch die Bedeutungsregeln der Sprache festgelegt ist. Analytisch wahre Sätze sind also solche, deren Wahrheit man schon allein aufgrund sprachlicher Regeln wissen kann, ohne auf die Welt "da draußen" zu sehen. (Analog gibt es auch analytisch falsche Sätze).

Analytische Sätze zerfallen in logische und definitiv. Logisch determinierte Sätze sind solche, deren Wahrheitswert aus den Regeln der Logik folgt. Definitivische Sätze nennt man auch "extralogisch analytische" Sätze, bei ihnen wird der Wahrheitswert durch sprachliche Definitionen oder Bedeutungskonventionen festgelegt.

2. *Synthetische Sätze* sind solche, die nicht analytisch sind, d.h. solche, die etwas Inhaltliches bzw. Gegenständliches behaupten. Ihr Wahrheitswert ist nicht durch die

Regeln der Logik und der Sprache festgelegt. Hier können wir wieder unterscheiden zwischen *deskriptiven* Sätzen, die etwas über die *reale Welt* aussagen, über das, was IST, und *präskriptiven* oder auch ethischen Sätzen, die etwas darüber aussagen, wie die Welt sein SOLL, bzw. wie sie sein müßte wenn sie GUT wäre. Es gibt auch *gemischte* Sätze, die sowohl deskriptive wie präskriptive Bestandteile enthalten.

Deskriptive Sätze sind jene Satzart, mit der wir uns in der Wissenschaftstheorie fast ausschließlich beschäftigen werden - ihre Wahrheit herauszufinden ist ja Zielbestimmung der Wissenschaft. Deskriptive Sätze zerfallen in empirische Sätze und theoretische Sätze. Empirische Sätze sind solche, die keine theoretischen Begriffe enthalten, also nur empirische Begriffe oder logische Begriffe. Theoretische Sätze sind solche, die *auch* theoretische Begriffe enthalten.

Präskriptive bzw. ethische Sätze kann man wieder in Norm- und Wertaussagen einteilen. Mit ihnen befaßt sich die *Ethik* - die Wissenschaftstheorie befaßt sich mit ihnen nur am Rande, insofern es für die Wertneutralitätsthese oder die Abgrenzung empirischer Wissenschaft von Ethik bedeutsam ist.

Die *genaue Definition* von logisch wahren Sätzen, definitorisch wahren Sätzen, von deskriptiven versus präskriptiven Sätzen, geben wir in den folgenden Abschnitten. Mit der Unterscheidung zwischen empirischen und theoretischen Sätzen befassen wir uns in Kap. 4.

Die Unterscheidung zwischen analytischen und synthetischen Sätzen "Urteilen" geht auf Kant zurück. William v. O. Quine (und andere) haben argumentiert, daß die Unterscheidung zwischen analytischen und synthetischen Sätzen oft graduell oder vage, manchmal sogar willkürlich sei. Dies trifft, wie wir in Kap. 4 sehen werden, auf wissenschaftliche Theorien in gewisser Hinsicht zu. In anderen Bereichen ist die Unterscheidung jedoch unproblematisch. Heuristisch ist sie jedenfalls unentbehrlich und wertvoll.

2.2.2 Logik und logische Wahrheit

Grundlegend für die moderne Definition von logischer Wahrheit ist die Unterscheidung zwischen logischen und nichtlogischen (inhaltlichen) Begriffen. Logische Begriffe, wie "und", "oder" etc., bezeichnen keine Dinge oder Merkmale der Welt, sondern sagen sozusagen nur etwas über die mögliche Form der Welt aus (diese Charakterisierung stammt von Wittgenstein). Logische Begriffe sind "rein strukturell". Nichtlogische Begriffe, prototypisch deskriptive Begriffe, Eigennamen oder Prädikate, bezeichnen gewisse Dinge oder Merkmale der Welt. Das trifft indirekt auch auf präskriptive Begriffe wie "ist gut" oder "ist gesollt" zu - hier wird ein Bezug zwischen der realen und einer idealen Welt hergestellt.

Ein Satz soll nur logisch wahr heißen, wenn seine Wahrheit allein aus der

Bedeutung der in ihm enthaltenen logischen Begriffe folgt. Man sagt auch - wenn seine Wahrheit schon allein aus seiner *logischen Form* folgt. Hierzu müssen wir zuerst die *logische Form* eines Satzes definieren. Man erhält sie, wenn man alle nichtlogischen Begriffe eines Satzes durch Variablen ersetzt, sodaß nur mehr das logische "Skelett" übrigbleibt.

Logische Form eines Satzes: Ersetze alle nichtlogischen Begriffe durch Variablen (Platzhalter)

Die logische Form von "wenn Peter krank ist, dann ist seine Mutter traurig", ist "wenn p, dann q" -- denn der logische Begriff dieses Satzes ist das "wenn-dann", während die logisch unzerlegbaren Sätze "Peter ist krank" und "seine Mutter ist traurig" durch die Variablen p und q ersetzt werden.

Nun können wir die logische Wahrheit eines Satzes definieren als Eigenschaft seiner logischen Form:

Ein Satz ist logisch wahr genau dann, wenn jeder Satz, der dieselbe logische Form hat, wahr ist.

Analog heißt ein Satz logisch falsch bzw. widersprüchlich, wenn jeder Satz derselben logischen Form falsch ist.

BEISPIEL eines logisch wahren Satzes:

Wenn alle Menschen sterblich sind, dann gibt es keinen Menschen, der unsterblich ist.

Logische Form: Wenn alle F G sind, dann gibt es kein F, das G ist.

Formalisierung: $\forall x(Fx \rightarrow Gx) \rightarrow \neg \exists x(Fx \wedge \neg Gx)$

-- Jeder Satz dieser logischen Form ist logisch wahr.

Im Gegensatz dazu sind Sätze der Form Alle F sind G nicht logisch determiniert Sie können entweder definatorisch oder synthetisch determiniert sein.

Analog definiert man die logische Gültigkeit eines Schlusses. Generell gesprochen besteht ein *Argument* bzw. ein *Schluß* aus mehreren Sätzen, genannt *Prämissen*, und aus einer *Konklusion*, von der man will, daß sie aus den Prämissen "folgt". Wenn der Schluß *logisch* gültig ist, so heißt daß, daß mit Sicherheit die Konklusion wahr ist, wenn wirklich alle Prämissen wahr sind. Man sagt auch, es sei bei einem logischen Schluß *denknotwendig*, daß durch die Wahrheit der Prämissen auch die Wahrheit der Konklusion feststeht. "Denknotwendig" ist aber ein ungenauer Begriff. In der

modernen Logik definiert man die logische Gültigkeit wiederum über die logische Form eines Schlusses.

Ein Schluß ist logisch gültig genau dann, wenn für jeder Schluß, der dieselbe logische Form hat, folgendes gilt:

wannimmer alle Prämissen wahr sind, ist auch die Konklusion wahr.

BEISPIEL: Alle Metalle leiten Strom (Prämisse)
Dein Koffer ist metallisch . (Prämisse)
 Daher leitet dein Koffer Strom (Konklusion)

Logische Form: Alle F sind G Formalisierung: $\forall x(Fx \rightarrow Gx)$
Dieses a ist ein F Fa
 Daher ist dieses a ein G Ga

Der Schluß ist in der Tat gültig. D.h., *jeder Schluß* dieser logischen Form muß gültig sein.

Ungültiger Schluß (Beispiel): Alle F sind G
Dieses a ist ein G
 Daher ist dieses a ein F

Daß die logischen Gesetze von inhaltlich-realtätsbezogenen Überlegungen unabhängig sind, muß so sein, denn nur dadurch kann die Logik beanspruchen, jene Gesetze auszudrücken, die denotwendig sind, also von der faktischen Beschaffenheit der Wirklichkeit unabhängig sind. - Insbesondere ist die Gültigkeit eines Schlusses daher unabhängig davon, ob seine Prämissen faktisch wahr sind (!! - das wird immer wieder mißverstanden). Wenn ein Schluß gültig ist, so ist er es also auch, wenn seine Prämissen falsch sind.

2.2.3 Bedeutungskonventionen und definitiorische Wahrheit

Definitionen wie "Ein Junggeselle ist ein unverheirateter Mann" scheinen ebenfalls 'denotwendig' zu sein, dennoch handelt es sich hier um etwas anderes als um logische Wahrheiten. Bei Definitionen bzw. Bedeutungskonventionen geht es um die Bedeutung *nichtlogischer* Begriffe. "Junggeselle", "unverheiratet", "Mann" sind nicht logische Begriffe. In Bedeutungskonventionen wird die Bedeutung solcher nichtlogischen Begriffe durch *Konvention* festgelegt. Definitiorisch wahre Sätze sind

also zwar ebenfalls unabhängig von der realen Welt allein durch sprachliche Regeln wahr - und daher analytisch wahr - bei ihnen hängt ihre Wahrheit jedoch nicht allein von der Bedeutung logischer Begriffe ab, sondern sie hängt von der spezifischen Bedeutung nichtlogischer Begriffe wie "Junggeselle" usw. ab. M.a.W., die Wahrheit definitorisch wahrer Sätze ist nicht allein eine Sache ihrer logischen Form - wohl aber allein eine Sache der Bedeutungsregeln der zugrundeliegenden Sprache. (Diese Auffassung von definitorisch wahren Sätzen geht auf Carnap zurück.)

Ein Satz ist definitorisch wahr genau dann wenn seine Wahrheit auf gewissen Bedeutungskonventionen für nichtlogische Begriffe beruht, welche in der zugrundeliegenden Sprache respektive Sprechergemeinschaft verankert sind.

BEISPIEL: Alle Junggesellen sind unverheiratet

Alle Schimmel sind weiß (Logische Form: Alle S sind W)

Die Länge des Urmeters in Paris ist ein Meter.

Es wäre absurd, wollte man die Wahrheit des Satzes "Junggesellen sind unverheiratet" durch empirische Untersuchung von Junggesellen herausfinden wollen - so haben wir die Bedeutung von "Junggeselle" eben festgelegt.

Allerdings trifft das Quinesche Argument hier doch gelegentlich zu - oft ist es bei einer gegebenen natürlichen Sprache nicht so eindeutig, was als Bedeutungskonvention und was als empirische Aussage gelten soll. Z.B. sind Schimmel per Konvention Pferde, die weiß sind - aber was ist mit Grauschimmeln?

Sätze, die weder logisch noch definitorisch wahr sind, sind synthetisch wahr - ihre Wahrheit kann man nur durch empirische Untersuchungen herausfinden.

Beispiel von synthetischen Sätzen: Alle Raben sind schwarz

Die Länge deiner Hose ist ein Meter.

Definitorisch wahre Sätze zerfallen drei Gruppen:

a) *Explizitdefinitionen*, bei denen ein Begriff (das Definiendum) durch einen Komplex von anderen Begriffen explizit definiert wird

b) *Bedeutungspostulate*, die nicht die Form einer Explizitdefinition haben, aber ebenfalls "axiomatisch" festgelegt sind

Gruppen a und b kann man auch definitorische *Axiome* nennen

c) schließlich solche, deren Wahrheit indirekt aus einigen definitorischen Axiomen logisch folgt.

(Definiendum)

(Definiens)

Explizitdefinition: $\forall x: x \text{ ist ein Schimmel} \leftrightarrow x \text{ ist ein Pferd und } x \text{ ist weiß}$

Andere definitorisch wahre Sätze:

Bedeutungspostulate:

Wenn etwas rund ist, dann hat es keine Ecken

Abgeleitete definitorisch wahre Sätze:

Wenn eine Person nicht männlich ist, dann ist sie kein Junggeselle.

Eine wichtige Anforderung an Explizitdefinitionen ist *Nichtzirkularität* - der Definiendumbegriff darf im Definiens nicht vorkommen. Es gibt aber auch sogenannte *rekursive* Definitionen, die in der Logik besprochen werden, wo dies etwas anderes ist. -- Die wichtigste Forderung ist aber:

FORDERUNG AN DEFINITORISCH WAHRE SÄTZE:

Sie müssen, relativ zum bereits akzeptierten Hintergrundwissen, *empirisch nichtkreativ sein*, d.h. ohne neuen Tatsachengehalt sein.

Ergo: Kein Begriff darf zweimal in verschiedener Weise definiert werden.

Z.B.: 1 Meter = Die Länge des Urmeters in Paris (1)

1 Meter = Die Länge eines Pendels in Seehöhe mit
1 sec Schwingungsdauer (2)

(1) + (2) implizieren folgenden synthetischen (empirischen) Satz: Die Länge des Urmeters in Paris = die Länge eines Pendels in Seehöhe mit 1 sec Schwingungsdauer.

Ergo: man darf nicht (1) und (2) *zugleich* als Definition der Längeneinheit 1m ansehen! Ist (1) die geltende Definition, so ist (2) eine empirische Hypothese; ist umgekehrt (2) die geltende Definition, so ist (1) eine empirische Hypothese.

Die in wissenschaftlichen Lehrbüchern zu findenden 'Definitionen', wie z.B. "Kraft = Masse mal Beschleunigung", oder "Volksschule - Institution zur Vermittlung von Grundwissen für 6-10jährige" erfüllen diese Forderung häufig *nicht*; d.h. sie sind mehr als bloße Begriffskonventionen und haben impliziten Tatsachengehalt. In noch stärkerem Maß gilt dies für empirische Operationalisierungen theoretischer Terme (die fälschlicherweise oft als 'operationale Definitionen' bezeichnet wurden). Dies entdeckt zu haben, war eine wichtige Errungenschaft der Wissenschaftstheorie (im Anschluß an Carnap), auf die wir noch zurückkommen. Wie wir sehen werden, ist vieles in wissenschaftlichen Theorien, was oft lax als 'Definition' bezeichnet wird, ein indirekt empirisch gehaltvoller Satz, der somit indirekt empirisch zu bewähren ist .

2.2.4 Deskriptive versus präskriptive Sätze

Wir nennen Normoperatoren, Wertprädikate und Wertoperatoren zusammenfassend präskriptive Begriffe. - Die Schwierigkeit bei der intendierten Unterscheidung liegt darin: nicht jeder Satz, der einen präskriptiven Begriff enthält, ist selber präskriptiv(!). Z.B. sind die Sätze

Peter glaubt, daß der Stehlen schlecht ist

Peter glaubt, daß man nicht stehlen soll

rein empirisch-deskriptiv - sie sagen etwas über die faktische Wert- bzw. Normeinstellung Peters aus. Dasselbe gilt für alle Wert- und Normaussagen, die hinter Phrasen wie "Person X glaubt, denkt, fühlt, sagt, will..." stehen. All das sind deskriptive Aussagen. Dies ist für die Beurteilung der Norm-oder Wertfreiheit gewisser Wissenschaftsdisziplinen sehr wichtig: Disziplinen wie die Geschichtswissenschaft, die Soziologie, etc., die faktische Wert- und Normeinstellungen von Personen untersuchen, machen deshalb noch lange nicht selber präskriptive Behauptungen.

Logisch kann man sagen, daß ein präskriptives Symbol in einem Satz genau dann *neutralisiert* wird, wenn er im Bereich eines sogenannten *neutralisierenden deskriptiven Operators* liegt. Das sind epistemische Operatoren, volitive Operatoren, Sprechaktoperatoren - aber auch alle anderen denkbaren Operatoren, die sich auf innere Zustände von Personen beziehen. Der Bereich eines Operators ist ganz einfach jene Teilformel, auf die der Operator sich bezieht. Damit können wir definieren:

Ein Satz ist rein deskriptiv g.d.w. er entweder kein präskriptives Symbol enthält oder aber jedes präskriptive Symbol, das er enthält, im Bereich eines neutralisierenden deskriptiven Operators liegt.

Unter den nicht rein deskriptiven Sätzen können wir zwei Arten unterscheiden: rein präskriptive Sätze und gemischte Sätze. Dabei kommt es auf die Teilformeln der Sätze an - jede Satzteil, die ihrerseits grammatisch wohlgeformte Satzformeln sind:

Ein Satz ist rein präskriptiv g.d.w. jede seiner deskriptiven Teilsätze im Bereich eines Wert-oder Normoperators liegt.

(Man überlege sich, warum in diesem Fall automatisch jeder deskriptive Teilsatz im Bereich eines Wert- oder Normoperators liegen muß, der seinerseits nicht im Bereich eines neutralisierenden deskriptiven Operators liegt.)

Ein weder rein deskriptiver noch rein präskriptiver Satz heißt gemischt.

BEISPIELE:

Deskriptive neutralisierenden Operatoren bezeichnen wir mit Großbuchstaben, präskriptive Operatoren mit Fettbuchstaben. Teilsätze unterstreichen wir.

Peter glaubt, daß Stehlen schlecht ist.

GLAUBT(Peter, **Schlecht**(Stehlen)) - rein deskriptiv

Stehlen ist verboten.

Verboten(Stehlen) - rein präskriptiv

Für jemand der Hunger leidet, ist Stehlen erlaubt.

$\forall x(\text{Wenn } \underline{x \text{ leidet Hunger}}, \text{ dann } \underline{\text{Erlaubt}}(x, \text{Stehlen}))$ - gemischt.

Peter glaubt, daß Stehlen schlecht ist, obwohl er selbst stiehlt.

GLAUBT(Peter, **Schlecht**(Stehlen)) und Stiehlt(Peter). --- rein deskriptiv

Wenn Stehlen erlaubt ist, dann gibt es kein Recht auf Eigentum.

Wenn Erlaubt(Stehlen), denn nicht: Recht(Eigentum). - rein präskriptiv

Eine sehr wichtige Art gemischter Sätze sind die sogenannten *bedingten Normsätze bzw. Wertsätze* "wenn das und das der Fall ist, dann soll das und das der Fall sein". Z.B. ist der Satz "Alle sich untereinander in Konkurrenz befindlichen Menschen sollen sich achten" ($\forall x \forall y (Mx \wedge My \wedge Kxy \rightarrow S(Axy))$) eine bedingte Norm.

2.2.5 Klassifikation von Satzarten nach ihrem Allgemeinheitsgrad

Die wissenschaftstheoretisch wichtigste Klassifikation von Satzarten ist die nach ihrem Allgemeinheitsgrad, zu der wir jetzt kommen.

Deterministische Allsätze

Für alle x: wenn x ist A, dann x ist K ($\forall x(Ax \rightarrow Kx)$)

A = *Antezedens*, K = *Konsequens*

rein
(raumzeitlich unbeschränkt)

raumzeitlich
beschränkt

Generelle Sätze

Statistische Sätze

$r\%$ aller A's sind K's $(p(K/A) = r)$

rein (raumzeitlich unbeschränkt)	raumzeitlich beschränkt
-------------------------------------	----------------------------

Singuläre Sätze: dieses a ist ein A, und es ist ein (kein) K.

Existenzsätze: Es gibt (irgendwo) ein A, das ein (kein) K ist.

(Gemischte Sätze, z.B. Allexistenzsätze, oder Existenzallsätze).

Erläuterungen:

Generelle Sätze sind im allgemeinen das, was in der Wissenschaft als Kandidat für eine *Gesetzesaussage* angesehen wird. Die Grenzen zwischen Gesetzesaussagen und singulären Einzelfallaussagen sind aber, wie noch zu sehen, fließend.

1. *Deterministische Allsätze*: Die angegebene Form einer Allimplikation ist nur die *einfachste Form* eines solchen Satzes. Im allgemeineren Fall handelt es sich um Sätze, die mit einem oder mehreren Allquantoren beginnen, auf die eine Formel folgt, welche man auch Matrix des Satzes nennt. In so gut wie allen wissenschaftlichen Allsätzen hat die Matrix die Form einer Implikation hat (Allimplikation, Implikationsgesetz : $\forall x(Fx \rightarrow Gx)$), oder ist auf Formen zurückführbar ist, die auf der Implikation aufbauen, z.B. die Form einer Äquivalenz (Äquivalenzgesetz, $\forall x(Fx \leftrightarrow Gx)$), gleichbedeutend mit $\forall x((Fx \rightarrow Gx) \wedge (Gx \rightarrow Fx))$), oder, bei quantitativen Gesetzen, die Form einer numerischen Identität zwischen quantitativen Größen ($\forall x(f(x)=g(x))$ - gleichbedeutend mit $\forall x \forall r(f(x)=r \leftrightarrow g(x)=r)$, r ist Zahlenwert).

Eine wichtige Unterklasse ist die der reinen raumzeitlich unbeschränkten Allsätzen oder Universalsätzen -- solche, wo in der Matrix keine Individuenkonstanten vorkommen, also insbesondere keine Raumzeitbeschränkungen auftreten. Inhaltlich gesehen ist also ein reiner deterministischer Allsatz einer, der über potentiell alle Objekte des Universums etwas aussagt, eine Prognose liefert, also wo weder der Antecedensbereich noch der Konsequenzbereich irgendeiner singulären bzw. individuellen Einschränkung unterworfen ist.

Reine generelle deterministische Sätze sind Kandidaten für *echte* Naturgesetze bzw. Kausalgesetze.

Beispiele:

- (1) Alle Körper ziehen sich gegenseitig an
 Für alle x und y: wenn x ein Körper ist und y ein Körper ist, dann zieht
 x y an und y zieht x an.
- (2) Für alle idealen Gase x gilt: Druck(x) mal Volumen (x)
 = Gaskonstante mal Molzahl(x) mal absolute Temperatur(x)

Die zweite Unterklasse sind die raumzeitlich beschränkte Allsätze: Diese Universalsätze enthalten Individuenkonstanten, nämlich Raumzeitkonstanten, welche ihren raumzeitlichen Anwendungsbereich einschränken. Einige Beispiele:

- (1) Alle Körper nahe der Erdoberfläche fallen mit Beschleunigung g nach unten
 (2) Alle Säugetierarten in Polargebieten haben verglichen zu ihren Artgenossen in warmen Ländern eine rundliche Form.
 (3) Bis ungefähr 10000 v. C. lebten alle Menschen vom Jagen und Sammeln.
 (4) Im Mittelalter beruhte aller Grundbesitz und Landwirtschaft auf dem Lehensprinzip.
 (5) Alle Äpfel in diesem Korb sind rot.

All diese Allsätze haben also die Form $\forall x((st(x) \in k \wedge Fx) \rightarrow Gx)$, st=Funktor, der jedem x sein Raumzeitintervall zuordnet, irgendein bestimmtes Raumzeitintervall, \in Elementrelation, "st(x) \in k" besagt also "x liegt im Raumzeitintervall k". Man stellt hier also eine gewisse Regelmäßigkeit fest, die nur auf einen bestimmten Raumzeitbereich zutrifft. Raumzeitlich eingeschränkte Gesetze kommen typischerweise in den Geschichtswissenschaften, und den Human- und Sozialwissenschaften vor. Die Frage, ob hierbei *überhaupt noch von einem wissenschaftlichen Gesetz gesprochen werden kann*, hängt vorallem von der Größe des raumzeitlichen Anwendungsbereiches - also des Antezedensbereiches - ab. Dies sieht man schön an der Hierarchie der Beispiele (1) - (5). (1) ist eindeutig etwas, das man als Gesetz bezeichnen würde. (2) auch noch eindeutig - bei (3) und (4) ist die Sache nicht mehr so klar; schließlich (5) würde man - obwohl es auch obige logische Form hat - sicher nicht als Gesetzesaussage, sondern als Einzelfallaussage ansehen, die über diesen Korb spricht.

Die Frage, welche Arten von Allsätzen als Kandidaten für echte Naturgesetze in Frage kommen, wurde in der Wissenschaftstheorie viel diskutiert - man nennt das das Problem der *Gesetzesartigkeit* (z.B. Hempel, Goodman, E. Nagel, u.a.). Es wurde beispielsweise vorgeschlagen, raumzeitlich eingeschränkte Allsätze dann als

gesetzesartig anzusehen, wenn sie aus reinen fundamentalen Naturgesetzen plus singulären Anfangsbedingungen logisch ableitbar sind. Das ist z.B. bei unseren Sätzen (1) und (82) der Fall. Ernest Nagel hat aber gezeigt, daß derartiges auch bei "Pseudoallsätzen" vom Typ (5) der Fall sein kann -- sein Beispiel ist der Allsatz "Alle Nägel an Smith's Wagen sind rostig" -- der aus Gesetzen über Wetterbedingungen in Schottland und den singulären Fakten, daß Smith Schotte ist und sein Wagen immer im Freien steht, abgeleitet werden kann. Somit ist die Grenze zwischen gesetzesartigen Sätzen und nichtgesetzesartigen Sätzen im Falle raumzeitlich eingeschränkter Sätze tatsächlich ein gradueller und pragmatischer, von der Größe des Antezedensbereiches und damit von der prognostischen Verwertbarkeit eines solchen Satzes abhängig.

2. *Statistische Sätze* -- sie sind auch *genereller Natur*, zumindest wenn sie raumzeitlich unbeschränkt sind, also über eine potentiell unendliche Grundgesamtheit sprechen. Dennoch zögern wir, sie "Allsätze" zu nennen, weil sie das streng genommen nicht sind. Ein statistischer Satz "r% aller A sind K" sagt nichts über alle Individuen vom Typ A aus, sondern er sagt etwas über die Klasse der As aus, nämlich daß sie sich mit der Klasse K zu r% von A überlappt. Aus diesem Grund gibt es zwischen statistisch generellen Sätzen und Singulärsätzen keine zwingenden logischen Beziehungen (s.u.). Daraus resultiert z.B. der grundlegende Unterschied zwischen deterministisch-deduktiven und statistisch-induktiven Erklärungen - letztere sind hintergrundrelativ, erstere nicht - oder die Tatsache, daß deterministische Gesetze durch Einzelbeobachtungen streng falsifizierbar sind, statistische dagegen nicht (s.u.).

Reine statistisch generelle Sätze: Eine in diesem Fall wichtige Frage ist, ob es sich hier um einen

epistemischer Indeterminismus: Bloße Wahrscheinlichkeit statt Sicherheit wegen Unvollständigkeit unseres Wissens handelt, oder wie in der Quantenmechanik ein

objektiver Indeterminismus: Bloße Wahrscheinlichkeit statt Sicherheit der Naturgesetze als objektive Indeterminiertheitseigenschaft der Natur vorliegt. Objektiv indeterministische reine statistische Allsätze sind Kandidaten für echte statistische Naturgesetze.

Bei raumzeitlich beschränkten statistisch generellen Sätzen gelten dieselben Abstufungen wie oben bei den raumzeitlich beschränkten deterministischen Allsätzen.

BEISPIELE für generelle statistische Sätze:

- (1) 50% aller Cs¹³⁷-Atome (einer beliebigen Substanzmenge) sind nach 30 Jahren zerfallen
- rein, statistisches Naturgesetz, objektiver Indeterminismus
- (2) 80% aller Lungenkrebskranken waren schwere Raucher
- rein, epistemisch unvollständiges statistisches Gesetz
- (3) 70% aller Bettnässer-Kinder haben Eltern mit gestörter Beziehung
- rein
- (4) 95% aller Schweden sind protestantisch
-raumzeitlich beschränkt
- (5) 60% aller Äpfel in diesem Korb sind rot.
- nicht gesetzesartig

Es gibt aber auch schwächere statistische Aussagen als die der Form "die bedingte Wahrscheinlichkeit ist so und so hoch". Beispielsweise:

Qualitative statistische generelle Sätze:

Hochwahrscheinlich, Normisch:

Die meisten Bettnässer-Kinder haben Eltern mit gestörter Beziehung:

$$p(\text{Gest.Bez.}(x) / \text{Bettn.}(x)) = \text{hoch} \in [0.7, 0.9]$$

Normalerweise können Vögel fliegen.

Hier wird kein numerischer Wert genannt, die Wahrscheinlichkeit ist nur vage hoch, was man eventuell mit einer Intervallaussage (untere und obere Grenzen) ausdrücken kann. Diese Aussagen spielen in der sogenannten modernen "nichtmonotonen Logik" eine wichtige Rolle.

Komparativ, Korrelationsaussagen:

Bettnässer-Kinder haben eher Eltern mit gestörter Beziehung

$$p(\text{Gest.Bez.}(x) / \text{Bettn.}(x)) \text{ ist größer als } p(\text{Gest.Bez.}(x))$$

Signifikant viele Bettnässer-Kinder haben Eltern mit gestörter Beziehung

Differenz: $p(\text{Gest.Bez.}(x) / \text{Bettn.}(x))$ versus $p(\text{Gest.Bez.}(x))$ ist

"hinreichend" hoch (zu hoch, um bloßer Zufall zu sein).

In der Statistik nennt man das Konsequenzprädikat übrigens auch *Kriteriumsvariable*, das Antecedensprädikat *Prädiktorvariable*. Den Bereich aller Objekte, auf den sich der Wahrscheinlichkeitsfunktork p bezieht, nennt man in der Statistik auch Grundgesamtheit (oder Population). Die empirischen Gesetzesaussagen in der Psychologie, Ökonomie und Sozialwissenschaften sind typischerweise statistischer Natur. Es handelt sich bei statistischen Gesetzen um die schwächste Form von

Regelmäßigkeiten.

3. *Singulärsätze*: Das sind allgemein solche Sätze, die keine Quantoren enthalten (weder Existenz- noch Allquantoren), und die auch keine statistischen Häufigkeitssätze sind. Sie sagen etwas über eines oder mehrere Einzeldinge aus. Eine besondere Form singulärer Sätze sind unnegierte oder negierte Atomsätze, der Form Fa , oder $\neg Fa$, oder Rab , oder $\neg Rab$, etc. Die ist die logische Form von *Basissätzen* bzw. des weiteren von *Beobachtungssätzen*. zumindest nach Carnap und Hempel. Popper hatte ihnen eine etwas andere Form gegeben, raumzeitlich eng beschränkte Existenzsätze, aber diese Form hat Nachteile.

4. *Existenzsätze*: Dies ist die logisch schwächste Satzart. Existenzsätze sind Sätze, die mit Existenzquantoren beginnen, auf die eine Matrix folgt. Meistens hat die Matrix die Form einer Konjunktion, z.B. $\exists x(Fx \wedge Gx)$ - es gibt einen schwarzen Schwan, es gibt einen Impfstoff gegen Pest, usw.

5. *Gemischte Sätze*: Natürlich gibt es auch quantifizierte Sätze, die sowohl All- wie Existenzquantoren enthalten, sie sind sogar oft wissenschaftlich wichtige Arten von Gesetzen, z.B.

Jede Handlung hat ein Motiv: $\forall x \exists y (\text{Handlung}(x) \rightarrow \text{Motiv}(y,x))$, oder
 Jede Substanz hat einen Siedepunkt: $\forall x \exists r (\text{Substanz}(x) \rightarrow \text{Siedepunkt}(x) = r)$,
 $r =$ reelle Zahl.

Allgemein kann man jeden gemischt-quantifizierten Satz, der zumindest einen Allquantor enthält, als einen Kandidaten für ein wissenschaftliches Gesetz ansehen.

2.4.6 Deduktive Beziehungen zwischen den Satzarten geordnet nach Allgemeingrad

\hat{U} steht für die logische Folgebeziehung

Deterministischer genereller Satz \hat{U} Singulärer Satz
 (Alle A sind K) \hat{U} (wenn a ist A, dann a ist K)

Deterministischer genereller Satz + Singulärsatz \hat{U} Singulärsatz (Deduktiv-
 nomologisches Erklärungsschema)
 (Alle A sind K; und a ist A) \hat{U} a ist K

Singulärsatz falsifiziert deterministischen generellen Satz

Singulärsatz \hat{U} Negation eines deterministischen generellen Satzes
 a ist A und a ist nicht K \hat{U} nicht: Alle A sind K

Existenzsatz falsifiziert deterministischen generellen Satz
 Es gibt ein x, das A ist und nicht K ist \hat{U} nicht: Alle A sind K

Deterministischer genereller Satz ist logisch äquivalent mit der Negation eines Existenzsatzes
 Alle A sind K \hat{E} nicht: Es gibt ein A, das kein K ist

Singulärsatz \hat{U} Existenzsatz+
 a ist ein A \hat{U} es gibt ein x, das ein A ist.

Zwischen statistischen generellen Sätzen und statistischen singulären Sätzen gibt es keine logischen Beziehungen, sondern nur Wahrscheinlichkeitsbeziehungen

2.4.7 Klassifikation von Satzarten nach ihrem methodologischen Status

Basissätze: Darunter verstehen wir (a) unnegierte oder negierte Atomsätze, die zudem (b) empirische Sätze sind, also neben logischen nur empirische Begriffe enthalten. (a) ist die syntaktische Formbedingung, (b) die semantische Bedingung.

Beispiel: Dort fällt ein Stein zu Boden. Dieser Rabe ist schwarz.

Man erinnere sich auch der Unterscheidung zwischen aktualen und hypothetischen Basissätzen.

Unter *Beobachtungssätzen* wollen wir etwas toleranter alle *singulären empirischen Sätze* verstehen.

Empirische Gesetze: Das sind generelle Sätze (gemäß unserer Klassifikation), die hinreichend allgemein sind, um als Gesetze zu gelten, und die zudem empirische Sätze sind, also neben logischen nur empirische Begriffe enthalten.

Beispiel: Alle Gegenstände, die sich in Erdnähe in der Luft befinden und nicht festgemacht sind, fallen zu Boden. Oder: Alle Raben sind schwarz.

Theoretische Gesetze: Das sind generelle Sätze, die auch theoretische Begriffe enthalten können.

Beispiel: Kraft = Masse mal Beschleunigung. Schwarze Körper absorbieren Lichtwellen.

Theorien: Das sind Systeme von empirischen und theoretischen Gesetzen plus

theoretischen Existenzsätzen. Wir besprechen sie später.

Beispiele: Die Newtonische Physik. Die Wellentheorie des Lichts.

Unter einer *Hypothese* versteht man allgemein einen Satz, der noch nicht durch Beobachtung bestätigt ist. Einer Hypothese kann also jede Satzart im obigen Sinn sein.

Somit zeichnen sich Hypothesen nicht durch eine gemeinsame logische Struktur aus - es können Allsätze, Existenzsätze und Singulärsätze sein - sondern lediglich durch ihren gemeinsamen methodologischen Status.

2.5 Verifikation, Falsifikation, Bestätigung (Bewährung) und Schwächung

Eine Hypothese ist

verifizierbar, wenn sie aus einer endlichen Menge von Beobachtungssätzen logisch folgt

falsifizierbar, wenn ihre Negation aus einer endlichen Menge von Beobachtungssätzen logisch folgt

Eine Gesetzhypothese ist verifiziert, wenn sie aus einer Menge von *aktualen* Beobachtungssätzen logisch bewiesen wurde -- man weiß dann mit Sicherheit, daß sie wahr ist; gegeben daß man sich der Wahrheit der Beobachtungssätze sicher ist. -- Analog für die Falsifikation. Man spricht auch, um die Sicherheit zu betonen, von *definitiver* Verifikation bzw. Falsifikation.

Eine Hypothese ist

bestätigbar, wenn sie durch eine endlichen Menge von Beobachtungssätzen zumindest wahrscheinlich oder "plausibel" gemacht wird.

falsifizierbar, wenn ihre Negation durch eine endlichen Menge von Beobachtungssätzen zumindest wahrscheinlich oder "plausibel" gemacht wird.

Daß Gesetzhypothesen zumindest bestätigbar und schwächbar sein müssen, besagt nur, daß es prinzipiell möglich ist, sie einer empirischen Überprüfung zu unterziehen, die in ihrem Ausgang dann "eher für" oder "eher gegen" die Hypothese spricht (wobei dieses "für" oder "gegen" oft relativ ist zum Stand der Alternativhypothesen, s. das Kap. über Induktion in 1). Voraussetzung für Bestätigbarkeit und Schwächbarkeit ist also nur, daß die Hypothese *empirischen Gehalt* hat, *empirische Konsequenzen* hat, an der man sie überprüfen kann - dies ist in der Tat eine Grundanforderung an alle synthetischen Hypothesen empirischer Wissenschaft.

Dagegen ist, wie gleich zu sehen, nicht jede Gesetzhypothese definitiv falsifizierbar, und schon gar nicht definitiv verifizierbar.

Reine deterministische Gesetze sind nicht streng verifizierbar (weil sie potentiell unendlich viele Anwendungsfälle haben), aber streng falsifizierbar (weil sie durch ein Gegenbeispiel widerlegt werden).

Reine statistische Gesetze sind weder streng verifizierbar (aus demselben Grund) noch streng falsifizierbar (sie sind auch nicht nur eine endliche Menge von Basissätzen logisch widerlegbar).

Natürlich sind reine deterministische Gesetze und reine statistische Gesetze

bestätigbar und schwächbar.

Raumzeitlich beschränkte Gesetze sind, egal ob deterministisch oder statistisch, im Prinzip immer streng verifizierbar und streng falsifizierbar, weil sie nur endlich viele Anwendungsfälle enthalten, die man im Prinzip aller nacheinander durchprüfen kann. Wir sagen "im Prinzip", weil man in der Praxis bei großem Anwendungsbereich - z.B. "alle Europäer" - auch hier nicht streng verifizieren, d.h. alles durchüberprüfen, sondern bestätigen, also auf eine repräsentative Stichprobe zurückgreifen wird.

Theorien sind ebenfalls weder (streng) verifizierbar noch streng falsifizierbar, aber aus einem anderen Grund, der mit dem sogenannten Holismus der Theorienüberprüfung zusammenhängt, was wir in Kap. 4 erklären.

Popper hat die sogenannte *Asymmetriethese* vertreten. Sie besagt, daß Gesetze zwar nicht verifizierbar, aber falsifizierbar sind. Offenbar trifft seine These nur auf reine deterministische Gesetze zu.

Wir fassen zusammen:

	Raumzeitl. beschr. Gesetze	Unbeschr. determ. Gesetze	Unbeschr. statist. Gesetze	Theorien
Verifikation +		-	-	-
Falsifikation	+	+	-	-
Bestätigung +		+	+	+
Schwächung	+	+	+	+

2.6 Literaturhinweise:

Eine einfache Einführung in die Logik, wo zugleich die logische Formalisierung geübt wird: Savigny, Eike v. (1976): Grundkurs im wissenschaftlichen Definieren, dtv, München. Ein einfaches logisches Einführungsbuch ist auch: W. Essler, R. Martinez (1983): Grundzüge der Logik, Klostermann, Frankfurt. Sehr umfangreich: W. Kliemann/N. Müller: Logik und Mathematik für Sozialwissenschaftler, I, II, W. Fink, München 1973, 1975. Als Lehrbuch in die Statistik zu empfehlen: J. Bortz, Lehrbuch der Statistik, Springer, Berlin,, 2.Auflage1985. Zur erwähnten logische Wahrscheinlichkeitstheorie: R. Carnap, R.Jeffrey: Studies in Inductive Logic and Probability, California Press, Press 1971. Zur Einführung in die subjektive

Wahrscheinlichkeitstheorie vgl. das Kap. 1 von G. Kleiter: Bayes-Statistik, de Gruyter, New York 1981. Zur Meßtheorie (Metrisierungstheorie) s. B. Orth: Einführung in die Theorie des Messens, Kohlhammer, Stuttgart 1974. - In den in Kap. 1 genannten wissenschaftstheoretischen Werken findet sich natürlich auch vieles über den Stoff von Kap. 2.

3. Empirische Gesetze und ihre Überprüfung

Die erste wichtige Bedingung dafür, einen gesetzesartigen Satz als wissenschaftliches Gesetz zu akzeptieren, ist natürlich seine vermutliche *Wahrheit*. Beispielsweise ist die Gesetzhypothese

"Alle Substanzen haben einen Siedepunkt"

zwar gesetzesartig, ihrer Form nach, aber nicht wahr, denn viele Substanzen zersetzen sich, bevor ihre Temperatur so hoch wird, daß sie schmelzen oder gar sieden könnten.

Aber es gibt noch eine zweite wichtige und oft übersehene Bedingung für Gesetzhypothesen: ihre *Relevanz*.

3.1. Die Relevanzbedingung - deterministischer und statistischer Fall

3.2.1 Wir behandeln zuerst den Fall eines *deterministischen Gesetzes*. Betrachten wir die Beispiele:

Alle Männer, die Antibabypillen nehmen, werden nicht schwanger

$$\forall x(Mx \wedge Ax \rightarrow \neg Sx)$$

Tollkirschen, an Vollmondnächten um Mitternacht gepflückt, haben halluzinogene Wirkung

$$\forall x(Tx \wedge Mx \rightarrow Hx)$$

Beides sind sowohl universelle wie wahre Implikationen; erfüllen also die bisherigen Bedingungen; dennoch stimmt etwas nicht: jeweils ein Konjunktionsglied des Gesetzesantecedens ist irrelevant für das Zustandekommen der Wirkung: Männer werden schon aufgrund ihrer biologischen Natur als Männer nicht schwanger, auch dann nicht, wenn sie keine Antibabypillen nehmen; und Tollkirschen haben immer halluzinogene Wirkung, auch wenn sie zu einer anderen Zeit als zu Vollmondmitternächten gepflückt werden. Das zweite Beispiel zeigt übrigens, daß es charakteristisch für viele Mythologien ist, durchaus korrekten empirische Zusammenhänge mit naturwissenschaftlich gesehen irrelevante Beifügungen zu befrachten.

Wir können diese Irrelevanzbedingung folgendermaßen präzisieren: Betrachten wir ein deterministisches Gesetz mit mehreren, einfachheitshalber zwei, konjunktiven Antecedensfaktoren ('Ursachefaktoren'); es hat die Form

$$\forall x(A1x \wedge A2x \rightarrow Kx)$$

Für alle x: wenn x die Eigenschaft A1 und A2 hat, dann hat es auch Eigenschaft K

Dann darf keine konjunktive Antecedenskomponente $A1x$ oder $A2x$ überflüssig sein, d.h. das Gesetz darf nicht auch schon ohne sie wahr sein - es darf also weder schon $\forall x(A1x \rightarrow Kx)$ noch $\forall x(A2x \rightarrow Kx)$ wahr sein.

Mehr intuitiv: nicht nur muß das ganze Antecedens $A1x \wedge A2x$ zusammen eine hinreichende Bedingung für Kx sein, sondern zusätzlich muß jede Komponente Aix auch ein *notwendiger Bedingungsanteil* sein. Analog gilt das Gesagte für Gesetze mit einem Antecedens, das mehr als zwei konjunktive Antecedensfaktoren enthält.

Ein irrelevanter Antecedensfaktor in einem Gesetz ist weder prognostisch relevant - d.h. für Prognosezwecke nötig - noch kausal relevant - d.h. kommt als Ursachefaktor in Frage. Merke jedoch: *prognostische Relevanz ist nur eine notwendige, aber nicht hinreichende Bedingung für kausale Relevanz* - dazu später.

3.1.2. Wir betrachten nun den Fall des *statistischen Gesetzes*. Die Grundgesamtheit sei die Menge aller an einer mittelschweren Erkältung erkrankten Personen. Die Wahrscheinlichkeit, daß eine solche Person bei regelmäßiger Einnahme hoher Dosen von Vitamin C innerhalb einer Woche wieder gesundet, ist sehr hoch, etwa 95%; d.h. es gilt das statistische Gesetz

$$p(G/V)=0,95$$

Dies wurde vor einiger Zeit tatsächlich behauptet. Doch es stellte sich heraus, daß ganz unabhängig von der Einnahme hoher Dosen Vitamin C 95% aller Erkältungserkrankten innerhalb einer Woche gesunden. Der Faktor V (=Einnahme hoher Dosen Vitamin C) ist also statistisch irrelevant; es gilt

$$p(G/V)=p(G)=95\%.$$

Im statistischen Fall gilt also: die bloße Höhe der bedingten Wahrscheinlichkeitsbeziehung $p(K/A)$ zwischen Antecedensmerkmal A und Konsequenzmerkmal K genügt nicht für einen relevanten gesetzlichen Zusammenhang; es muß zusätzlich die Bedingung erfüllt sein, daß das Antecedensmerkmal die bedingte Wahrscheinlichkeit des Konsequenzmerkmals $p(K/A)$ gegenüber der sogenannten Ausgangswahrscheinlichkeit $p(K)$ verändert. Wahrscheinlichkeitserhöhende Antecedensmerkmale heißen *positiv statistisch relevant*; wahrscheinlichkeitssenkende *negativ statistisch relevant*. Als prognostisch oder kausal relevante Merkmale für K kommen natürlich nur positiv relevante in Frage; während negativ relevante Faktoren für das Gegenteil von K, also für $\neg K$ prognostisch oder kausal relevant sind.

Wenn zwischen ein Merkmal A für eine zweites K statistisch relevant ist, sagt man

auch, zwischen A und E besteht eine *Korrelation*. Bei qualitativen Begriffen ist die Wahrscheinlichkeitsdifferenz $p(K/A)-p(K)$ das einfachste Korrelationsmaß. Es gibt verschiedene Korrelationsmaße, sowohl für einfache qualitative (zweiwertige) Begriffe (z.B. schwanger vs. nicht-schwanger, Vitamin C-Einnahme oder nicht) wie für quantitative Begriffe (z.B. Intelligenzquotient, Einkommen), welche man in Statistik-Lehrbüchern findet. Der Wert der Korrelation zwischen zwei Merkmalen A und K, $\text{Korr}(K,A)$, liegt immer zwischen -1 und +1. Bei einem positiven statistischen Zusammenhang ist die Korrelation positiv, wir schreiben $\text{Korr}(A,K)=\text{pos.}$, andernfalls negativ, $\text{Korr}(A,K)=\text{neg.}$, wenn keine relevante Beziehung vorliegt, gilt $\text{Korr}(A,K)=\text{null}$. - Auch bei statistischen Gesetzen muß man sich darüber klar sein, daß das Vorliegen einer positiv relevanten statistischen Beziehung zwischen A und K, d.h. einer positiven Korrelation $\text{Korr}(K,A)$ kein hinreichender Grund dafür ist, daß A auch ein Kausalfaktor (Ursachefaktor) für K ist, d.h. daß zwischen A und K eine Kausalbeziehung besteht.

Bei einem statistischen Gesetz mit zwei oder mehreren Antecedensmerkmalen, z.B. $p(K/A1 \wedge A2)=95\%$, muß sowohl A1 wie A2 statistisch relevant sein - was man durch sogenannte *Partialkorrelationen* ausdrückt: $\text{Korr}(K,A1/A2)$ ist die Korrelation zwischen K und A1 bei festgehaltenem A2; für qualitative Begriffe z.B. die Differenz $p(K/A1 \wedge A2)-p(K/A2)$. Die Relevanzbedingung für statistische Gesetze besagt, daß für alle konjunktiven Antecedensfaktoren ihre Partialkorrelation mit K nicht-null sein muß.

Irrelevante statistische Beziehungen sind ein erstes Beispiel von praktisch (sowohl in der Wissenschaft wie in der populären Darstellung und Auswertung von Wissenschaft) vorkommenden Fehlschlüssen. Obiges Beispiel vom Vitamin C war real: man empfahl tatsächlich lange Zeit Vitamin C als Gegenmittel gegen Erkältungen, bis genauere Untersuchungen die Irrelevanz von Vitamin C als wahrscheinlich erwiesen. Wie Huff (1956, S. 22) berichtet, ermangeln viele medizinische Tests über die Effektivität von Medikamenten mangelnder Vergleiche mit Kontrollgruppen, in denen die Relevanz der statistischen Daten geprüft wird (s. Kap. 3.3): so testete man ein Serum gegen Kinderlähmung bei 100 Kindern eines Dorfes, in dem kurz nach der Impfung eine Kinderlähmungsvirus auftrat. Kein Kind erkrankte - überragender Erfolg, hieß es. Tatsächlich erkrankte aber auch keines der nichtgeimpften Kinder - der Kinderlähmungsvirus zeigt nämlich nur in seltenen Fällen Wirkung. Ein anderes Beispiel ist Uri Geller's Uhrenheilung durchs Fernsehen: nach seiner Aufforderung der Zuseher, ihre alten Uhren aus den Kästen zu holen und zu sehen, ob sie wieder gingen, riefen tatsächlich einige hundert Leute an: "wahrhaftig, die Uhr geht wieder". Tatsächlich ist die Wahrscheinlichkeit, daß von einer Million alten weggelegten Uhren durch Herumhantieren einige hundert wieder zu Laufen beginnen, auch ohne Gellers Heilungskunst gegeben.

3.2. Die empirische Überprüfung von Gesetzen auf vermutliche Wahrheit (Bestätigung vs. Falsifikation) und auf Relevanz

Zuerst prüft man ein empirisches Gesetz auf vermutliche Wahrheit bzw. Falschheit. Die Prüfung eines Gesetzes auf Relevanz natürlich nur dann sinnvoll ist, wenn es noch nicht falsifiziert (geschwächt) ist; hat man ein Gesetz als falsifiziert erkannt, so erübrigt sich die Frage nach der Relevanz des Gesetzes.

3.2.1. Der deterministische Fall: Die Methode der Übereinstimmung (Bestätigung und Falsifikation) und des Unterschieds

Die Überprüfung eines solchen Gesetzes besteht aus zwei simultanen Tätigkeiten, die nach dem Logiker und Philosophen John Stuart Mill (1806-1873) die Methode der Übereinstimmung und des Unterschieds genannt werden. Die Methode der Übereinstimmung dient der Überprüfung auf Wahrheit/Falschheit, die Methode des Unterschieds dient der Überprüfung auf Relevanz/Irrelevanz.

Wir betrachten die Bestätigung einer deterministischen Gesetzhypothese der einfachen Form mit zwei Antecedensfaktoren, beispielsweise

"Alle Festsubstanzen (=A1) dehnen (=K) sich bei Erwärmung (=A2) aus", kurz: $\forall x(A1x \wedge A2x \rightarrow Kx)$.

Für das ganze Bündel der Antecedensfaktoren, in unserem Fall A1x und A2x zusammen schreiben wir einfach Ax.

1. Methode der Übereinstimmung: Damit ist herauszufinden, ob das Gesetz bestätigt, d.h. vermutlich wahr oder falsch ist. Man untersucht möglichst viele Individuen a_1, \dots, a_n , auf die das ganze Gesetzesantecedens zutrifft, wo also Aa_i gilt. Eine Menge solcher Individuen, auf die das Antecedens A zutrifft, heißt *A-Stichprobe*. Man untersucht nun für die Individuen der A-Stichprobe, ob sie alle die Eigenschaft K des Gesetzeskonsequens besitzen. Tun sie dies alle - d.h. führt unsere Untersuchung für jedes untersuchte Individuum a_i der A-Stichprobe zum Basissatz $Aa_i \wedge Ka_i$ - so ist das Gesetz bestätigt.

'Bestätigt' heißt, wie schon angeführt, 'vermutlich wahr', aber niemals 'sicher wahr' oder 'verifiziert', denn es ist immer möglich, daß in Zukunft noch nicht untersuchte Fälle a_j auftreten, die das Gesetz falsifizieren. Genau das meinte Popper, wenn er sagte, daß all unser Gesetzeswissen aus prinzipiellen Gründen immer nur mehr oder weniger gut bestätigtes Vermutungswissen (konjekturales Wissen) ist,

jedoch nie sicheres Wissen. Existiert jedoch nur ein einziges Individuum a , für das unsere Beobachtung zum Basissatz $Aa \wedge \neg Ka$ führte, so ist das Gesetz streng falsifiziert - denn aus $Aa \wedge \neg Ka$ folgt logisch $\neg \forall x(Ax \rightarrow Kx)$, also die Negation des Gesetzes.

Repräsentative Stichprobe: Es kommt weniger darauf an, daß man quantitativ sehr viele A-Individuen untersucht, sondern vorallem, daß man Individuen unter möglichst stark *variierenden Begleitumständen* untersucht. Generell ist die wichtigste Bedingung an jede Stichprobe (ob im deterministischen oder statistischen Fall), daß sie repräsentativ ist; und im deterministischen Fall bedeutet Repräsentativität nichts anderes als möglichst stark variierende Begleitumstände. Dies entspricht übrigens dem, was Popper mit strenger Prüfung gemeint hat. In unserem Beispiel müssen wir also zur Überprüfung des Gesetzes nicht nur möglichst viele feste Substanzen, die erwärmt wurden, untersuchen (z.B. 1000 Eisenstücke - das wäre trivial), sondern vorallem möglichst viele verschiedene Arten von Festsubstanzen. Tatsächlich entdeckte man in der organischen Chemie Feststoffe (gummiartige synthetische Stoffe), die sich bei Erwärmung zusammenziehen: das ursprüngliche Gesetz $\forall x(Ax \rightarrow Kx)$ war damit falsifiziert; man kann höchstens Gesetze mit spezifischerem Antecedens wie "Alle Metalle dehnen sich bei Erwärmung aus" formulieren.

2. Methode des Unterschiedes: Um zu überprüfen, ob das Gesetz - falls bestätigt - auch relevant ist, d.h. ob tatsächlich alle Antecedensfaktoren relevant sind, muß man zusätzlich für jeden Antecedensfaktor A_i eine sogenannte *A_i -Kontrollstichprobe* von Individuen untersuchen, auf die *alle Antecedensfaktoren außer A_i* zutreffen. Nur dann, wenn in der A_i -Kontrollstichprobe *nicht jedes* Individuum die Eigenschaft K aufweist, haben unsere Beobachtungen gezeigt, daß der Antecedensfaktor tatsächlich relevant ist. In unserem Beispiel muß untersucht werden, ob die Erwärmungs-Ausdehnungshypothese nicht eventuell nicht nur für Festsubstanzen, sondern für Substanzen beliebiger Aggregatzustände gilt - sodaß der Faktor Festsubstanz für den Erwärmungs-Ausdehnungszusammenhang irrelevant wäre. D.h. wir untersuchen z.B. auch Flüssigkeiten, die erwärmt wurden, auf Ausdehnung. In der Tat aber wissen wir von einer Flüssigkeit, mit der wir sehr vertraut sind, nämlich Wasser, daß sie sich bei Erwärmung zwischen 0 und 4 °C zusammenzieht - somit ist die Relevanz des Faktors 'Festsubstanz' zunächst gewährleistet.

Ein gutes Beispiel für die simultane Anwendung von Übereinstimmungsmethode und Unterschiedsmethode ist die *Entdeckung der Ursache des Kindbettfiebers durch Semmelweis* (geschildert von Hempel 1974, S. 11 ff).

Ausgangssituation: Die Anzahl der an Kindbettfieber erkrankten Mütter in der 1. geburtshilflichen Abteilung des Krankenhauses, an dem Semmelweis tätig war, war

ungemein hoch. Was war die Ursache?

Erster Hypothesensatz: die schlechte Verpflegung, die Überbelegung der Betten, meteorologische Einflüsse sind Ursachen. Diese konnte Semmelweiß durch Vergleich mit der 2. geburtshilflichen Abteilung widerlegen - hier nämlich trat das Kindbettfieber nicht auf, obwohl die genannten Faktoren ebenfalls alle anwesend waren. - Widerlegung durch die Methode der Übereinstimmung.

Zweiter Hypothesensatz: Es muß offenbar ein Faktor die Ursache sein, der in der 2. Abteilung, wo Kindbettfieber nicht auftritt, nicht vorhanden ist. (Suche nach relevantem Faktor gemäß der Methode des Unterschieds; 2. Abteilung fungiert als Kontrollgruppe). Da kamen in Frage: in der 1. Abteilung wurden die Frauen auf dem Rücken entbunden, in der 2. auf der Seite; oder in der 1. Abteilung kamen die Priester erst später als in der 2. Abteilung. Semmelweiß ordnete an, in der 1. Abteilung dies genau gleich zu machen wie in der 2.; Resultat: das Kindbettfieber blieb genauso hoch. - Widerlegung der Relevanz durch die Methode des Unterschieds.

Dritte, zum Ziel führende Hypothese: in der 1. Abteilung waren an der Geburtshilfe Medizinstudenten beteiligt, die zuvor aus dem Autopsiesaal kamen und noch Leichensubstanz an den Händen hatten - dies war in der 2. Abteilung nicht so. War das der Grund? Die Anordnung von Semmelweiß an die Medizinstudenten, sich die Hände vorher in Chlorkalk zu waschen, hatte sofort eine Senkung der Kindbettfiebersrate auf normale Werte zur Folge. (Dadurch: Bestätigung der Relevanz durch die Methode des Unterschieds; und durch das Wissen, daß - solange das Kindbettfieber auftrat - die Studenten der 1. Abteilung aus dem Autopsiesaal kamen: Bestätigung der Wahrheit durch die Methode der Übereinstimmung).

Mit der Findung dieses empirischen Gesetzes zwei wichtige Innovationen verbunden: die Theorie der Krankheitsentstehung durch Infektionsträger - z.B. Leichensubstanz - was allgemein die Entstehung der bakteriologischen Medizin einleitete, und die Möglichkeit von Infektionsschutz durch Desinfizierung mithilfe chemischer Substanzen wie Chlorkalk.

Vorzüge des kontrollierten Experiments gegen über bloßer Beobachtung ("Feldbeobachtung"): Man sieht an diesen Beispielen sehr deutlich die großen Vorzüge des kontrollierten Experiments gegenüber bloßer Beobachtung. Bei bloßer Beobachtung sind wir darauf angewiesen, daß uns das beobachtete Geschehen zufällig möglichst variierende Situationen präsentiert, aufgrund derer wir herausfinden können, welche Faktoren relevant sind und welche nicht. Das ist häufig nicht der Fall. Im Experiment dagegen *erzeugen* wir gezielt genau solche empirischen Situationen, die hinsichtlich spezifischer Faktoren systematisch variieren und uns über unsere Gesetzhypothesen maximal effektiven Aufschluß geben. Hätte z.B. Semmelweiß die Vorgänge seiner Klinik immer nur beobachtet, er wäre wohl nie

(außer durch großen Zufall) auf den richtigen Faktor gekommen.

Aus diesem Grund ist es auch in Disziplinen mit eingeschränkter oder fehlender Experimentiermöglichkeit - Soziologie, Wirtschaftswissenschaft, Geschichtswissenschaft - so schwierig, die Relevanz von Faktoren herauszufinden (Beispiel: "was ist für die Entstehung von Faschismus wirklich relevant?"). Typische 'Paradigmen' in diesen Disziplinen kennzeichnen sich meist dadurch aus, daß sie - oft in nicht wirklich empirisch geprüfter, sondern theoretisch-apriorischer Weise - gewisse Faktoren von vornherein als relevant ansehen und andere als irrelevant ausschließen - z.B. das ökonomistische Paradigma (Marx), das funktionalistisch-integrative Paradigma (Parsons, Luhmann), das Konfliktparadigma (Dahrendorf), das individualistisch-psychologistische Paradigma (Opp) in der Soziologie. Solange sich eine solche Relevanz/Irrelevanz-Einteilung nicht auch selber einer empirischen Überprüfung stellt, ist es gerechtfertigt, das Paradigma als 'ideologisch beladen' zu bezeichnen (falls die empirische Überprüfung nicht möglich ist, müssen eben alle Faktoren als möglicherweise relevant zugegeben werden).

3.2.2 Der statistische Fall: die Methode der Konfidenzintervalle und der Signifikanztests

Im statistischen Fall ist die Situation analog, bis auf zwei Unterschiede: erstens gibt es hier auch keine strenge Falsifikation, und zweitens wird die Repräsentativität von Stichproben etwas anders definiert. Nehmen wir zunächst das Beispiel:

80% aller Lungenkrebskranken waren schwere Raucher, kurz
 $p(K/A)=80\%$ ($Ax=x$ ist Lungenkranker, $Kx = x$ war Raucher).

Dabei wollen wir uns angenommen auf Österreicher zwischen 1980 und 1988 beziehen. Zunächst fällt auf, daß das Antecedens hier "Lungenkrebskrank" und das Konsequenz "schwerer Raucher" ist, also die statistische Beziehung hier genau umgekehrt zur Kausalbeziehung verläuft - aber das soll uns hier nichts stören (wie gesagt besagen Korrelationen von sich aus nichts über die Kausalbeziehung). - Wieder gibt es zwei Schritte:

1. Repräsentative A-Stichprobe - Konfidenzintervallmethode: Analog zur obigen 'Methode der Übereinstimmung' müssen wir das Gesetz $p(K/A)=80\%$ auf vermutliche Wahrheit überprüfen, indem wir eine repräsentative A-Stichprobe von n , sagen wie 100, Lungenkrebskranken befragen, ob sie schwere Raucher waren.

Die *Repräsentativitätsforderung* besagt im statistischen Falle nicht bloß, wie im deterministischen Fall - wo ein Gegenbeispiel zur Falsifikation genügt - daß die

untersuchten Individuen hinsichtlich der Umgebungsfaktoren möglichst extrem variieren sollen. Sie besagt spezifischer, daß alle sonstigen relevanten Merkmale in der A-Stichprobe gleich verteilt sein müssen wie in der Menge aller A-Individuen in der Grundgesamtheit (welche selber eine - andere - Grundgesamtheit, die 'A-Grundgesamtheit' ist; d.h. "Grundgesamtheit" wird in der Statistik variabel verwendet). Denn Lungenkrebs hängt natürlich noch mit anderen Faktoren als Rauchen kausal zusammen - z.B. in welcher Gegend man wohnt, welchen Beruf man hat, etc. - und damit der in der Stichprobe ermittelte Prozentsatz mit dem der Grundgesamtheit einigermaßen übereinstimmt, müssen diese restlichen Kausalfaktoren in der Stichprobe ebenso verteilt sein wie in der Grundgesamtheit. Die einfachste Methode, die zu gewährleisten, ist die *Zufallsstichprobe*: alle Individuen der A-Grundgesamtheit (hier lungenkrebskranke Österreicher 1980-1988) müssen *dieselbe Chance* haben, in die Stichprobe zu gelangen. Andere Methoden sind etwa die geschichtete Stichprobe - dazu näheres in Statistik-Lehrbüchern.

Wie schließt man nun von der angenommen repräsentativen Stichprobenhäufigkeit auf die Grundgesamtheitshäufigkeit? Angenommen, von 100 Lungenkrebskranken in der Stichprobe waren 75 schwere Raucher - ist das nun eine Bestätigung oder Falsifikation unseres obigen Gesetzes? Sicher ist nicht zu erwarten, daß Stichprobenhäufigkeit mit Grundgesamtheitshäufigkeit *genau* übereinstimmt, das ist wegen der Zufallsschwankungen sogar extrem unwahrscheinlich. Eine strenge Falsifikation gibt es hier also nicht. Man bedient sich der Methode der Konfidenzintervalle (das ist das üblichste Verfahren). Gegeben das Gesetz $p(E/A)=80\%$ ist wahr, so kann man die Wahrscheinlichkeit berechnen, daß ein Stichprobenhäufigkeitsresultat in Abhängigkeit von der Stichprobengröße n in einem bestimmten Intervall liegt (über die Binominalverteilung; bei quantitativen Variablen über die Gauß-Verteilung, wobei die Schätzung der Streuung miteingeht). Z.B.: mit 95% Wahrscheinlichkeit (oder 95,5%, 99%) liegt die Häufigkeit von K einer $n=100$ A-Stichprobe bei gegebenem Gesetz $p(K/A)=80\%$ im Intervall von 78 bis 82. Man nennt den pragmatische gewählte Wahrscheinlichkeitswert von 95% den *Konfidenzkoeffizienten* (Konfidenz = Vertrauen in das Gesetz), das damit vorausgesagte 95%-wahrscheinliche Stichprobenintervall 78 bis 82 das *Konfidenzintervall*. Letzteres wird umso kleiner, je größer der Stichprobenumfang ist.

Konfidenzintervall =

jenes Intervall, in dem die Stichprobenhäufigkeit mit einer Wahrscheinlichkeit \geq Konfidenzkoeffizient liegt, gegeben die statistische Gesetzeshypothese ist wahr.

Wenn nun das tatsächlich ermittelte Stichprobenresultat innerhalb dieses Konfidenzintervalls liegt, so wird man die Gesetzeshypothese beibehalten, sie ist

dann *schwach statistisch bestätigt*.

Falls das Stichprobenresultat aber außerhalb des Konfidenzintervalls liegt, dann hat es eine Wahrscheinlichkeit kleiner 5% -- dann sieht man die Gesetzhypothese als *stark statistisch geschwächt* an. In unserem obigen Beispiel ist das der Fall.

Nur *schwach* statistisch bestätigt ist die Gesetzhypothese im positiven Fall deshalb, weil es viele statistische Alternativhypothesen gibt, die ebenfalls durch dieses Stichprobenresultat schwach statistisch bestätigt sind. Angenommen etwa, 79 von 100 Lungenkrebskranke waren Raucher, und das bei Konfidenzkoeffizient = 0,95 berechnete Konfidenzintervall ist 78 - 82. Dadurch ist nicht nur die Gesetzhypothese, daß 80% schwere Raucher waren, schwach bestätigt, sondern alle Gesetzhypothesen, die um 2% um den Wert 79 herum liegen, also alle Gesetzhypothesen zwischen 77% und 81% -- denn für alle diese Gesetzhypothesen liegt der tatsächliche Stichprobenwert 79 von 100 in dem von ihnen mit 95% Wahrscheinlichkeit vorausgesagtem Stichprobenintervall (bei 77% z.B. 75 - 79; bei 78% z.B. 76 - 100, ..., bei 81% dann 79 - 83). Nur das statistische *Intervallgesetz*

Zwischen 77 und 81 % aller Lungen krebskranken waren schwere Raucher

ist durch das Stichprobenresultat 79 von 100 (bei Konfidenzkoeffizient 95%) *stark* statistisch bestätigt.

Oft liegt in der Statistik nicht der eben geschilderte Fall vor (Bestätigungs-/Falsifikationsfall), wo die Gesetzhypothese zuerst gegeben ist, und sie dann in einer Stichprobe auf Wahrheit überprüft werden soll, sondern der Fall des *induktiven Schlusses*, wo zuerst ein Stichprobenresultat vorliegt, von dem aus dann eine Gesetzhypothese aufgestellt wird. Hier geht man so vor: ausgehend vom Stichprobenresultat kann man das Konfidenzintervall berechnen, in dem die Grundgesamtheitshäufigkeit liegen muß, um mit 95% Wahrscheinlichkeit genau das erhaltene Stichprobenergebnis zu liefern, Man kommt damit zu statistische Intervallhypothesen über die Grundgesamtheit wie im obigen Beispiel, und diese sind dann statistisch stark bestätigt (näheres zu wieder in Statistik-Lehrbüchern).

Es ist zu betonen, daß der Konfidenzkoeffizient von ca. 95% zwar pragmatisch, aber nicht willkürlich gewählt ist. Wählt man ihn zu groß, z.B. 99,5 %, so wird das Konfidenzintervall zu breit, und zu wenig Hypothesen werden ausgeschlossen. Wählt man ihn zu klein, z.B. 75%, so wird das Konfidenzintervall zu klein, und man die Schwächung ist im negativen Fall nicht stark genug.

Wichtig ist auch, daß bei gegebenem Konfidenzkoeffizient das berechnete

Konfidenzintervall bei größerem Stichprobenumfang kleiner wird. Allerdings ist der Zusammenhang umgekehrt quadratisch - Konfidenzintervall ist proportional 1 durch das Quadrat des Stichprobenumfangs. Das heißt, macht man bei $p(K/A) = 80\%$ die Stichprobe größer, so wird die Wahrscheinlichkeit zufälliger Abweichungen der Stichprobenhäufigkeit von 0,8 immer kleiner; aber es gilt das Gesetz des abnehmenden Ertrags. Daher bringen Stichproben größer 100 nicht mehr viel. Stichproben unter 20 sind andererseits zu klein; hier werden die Zufallsschwankungen zu groß.

2. Kontrollgruppe - signifikante Korrelation: Um zu prüfen, ob das Merkmal A für K tatsächlich auch relevant ist, muß - analog zur Methode des Unterschieds - die Wahrscheinlichkeit von K in einer A-Stichprobe (Merkmalsstichprobe) mit der Wahrscheinlichkeit von K in einer repräsentativen A-Kontrollstichprobe (sogenannte Kontrollgruppe) beliebiger Österreicher verglichen werden. Angenommen in der A-Stichprobe sind von 100 80 schwere Raucher, in der Kontrollgruppe 60 - heißt dies, daß Rauchen die Wahrscheinlichkeit von Lungenkrebs erhöht, oder könnte die Stichprobenabweichung zwischen A-Stichprobe und A-Kontrollstichprobe nur *zufällig* bedingt gewesen sein? Wieder bedient man sich hier einer analogen Intervallmethode wie oben. Man kann die Wahrscheinlichkeit berechnen, daß die Abweichung zwischen A-Stichprobe und A-Kontrollstichprobe zufällig bedingt war. Man nennt diese die *signifikante Stichprobendifferenz*, und den entsprechend pragmatisch aber nicht willkürlich gewählten Koeffizienten den *Signifikanzkoeffizient*. Der Signifikanzkoeffizient wird meist mit 5% festgelegt (er ist also klein). Beträgt diese Wahrscheinlichkeit weniger als dieser Signifikanzkoeffizient, so nimmt man an, der untersuchte Faktor A ist tatsächlich für das Konsequenzmerkmal K statistisch relevant - man sagt, die Korrelation ist signifikant - andernfalls nimmt man an, daß die Abweichung zwischen A-Stichprobe und A-Kontrollstichprobe zufällig bedingt war und das A für K in Wirklichkeit nicht relevant ist.

Signifikante Stichprobendifferenz =

jener Betrag, den die Differenz zwischen A-Stichprobenhäufigkeit und A-Kontrollstichprobenhäufigkeit mit einer Wahrscheinlichkeit \leq Signifikanzkoeffizient übersteigt,

gegeben es besteht in der Grundgesamtheit kein Zusammenhang.

Angenommen in unserem Beispiel ist bei einem gegebenen Signifikanzkoeffizient von 5% die errechnete signifikante Stichprobendifferenz 10 von 100. Beträgt der Unterschied von A-Stichprobe und A-Kontrollstichprobe mehr als 10 - in unserem

Beispiel 20 - so ist die Hypothese, daß der Faktor A für den Faktor K relevant ist, statistisch stark bestätigt. Man sagt auch, es liegt eine *signifikante Korrelation* vor.

Falls die Differenz positiv ist, d.h. A zu einer Erhöhung von K führt, spricht man von einer *positiven* Korrelation, falls die Differenz negativ ist, von einer *negativen* Korrelation.

Falls die tatsächliche Differenz zwischen den beiden Stichproben *kleiner* als die signifikante Differenz ist, so ist die Hypothese, daß A für K relevant ist, *statistisch geschwächt*.

Dabei wird die Wahrscheinlichkeit der Zufallsbedingtheit einer gegebenen Abweichung wieder (umgekehrt quadratisch) kleiner, je größer die Stichproben sind. -- Die Annahme, daß die Beziehung 'irrelevant' ist, d.h. Merkmalsgruppe und Kontrollgruppe derselben Grundgesamtheit entstammen, nennt man in der Statistik übrigens auch "*Null-Hypothese*", die Relevanzannahme "*Alternativhypothese*".

Eine wichtige Voraussetzung beim Kontrollgruppenvergleich ist die *Vergleichbarkeit* von Merkmalsgruppe und Kontrollgruppe hinsichtlich der restlichen wirksamen Faktoren: diese müssen in beiden Gruppen gleich verteilt sein, damit man aus einer gegebenen signifikanten Häufigkeitsdifferenz berechtigt schließen kann, daß tatsächlich A für die ermittelte Häufigkeitsdifferenz relevant war. Diese Vergleichbarkeit kann man am ehesten noch im wissenschaftlichen Experiment, wo das Antecedensmerkmal A neu erzeugt wird, erreichen (durch sogenannte Randomisierung: auf eine repräsentative Stichprobe der Grundgesamtheit wird das Merkmal A zufällig verteilt, danach wird die Stichprobe in A-Merkmalsgruppe und Kontrollgruppe aufgeteilt). Bei Untersuchungen bereits vorhandener Merkmale durch Umfragen oder Tests ('Quasiexperimente') sind Verzerrungen in der Vergleichbarkeit unvermeidlich - sie sind hier für die "Scheinkorrelationen", die wir im nächsten Abschnitt besprechen.

3.2.3 Mögliche Fehlerquellen

Fassen wir zusammen: Bei der Überprüfung eines statistischen Gesetzes $\text{Korr}(K,A) = r$ gehen insgesamt drei Voraussetzungen ein: 1) die A-Merkmalsstichprobe muß repräsentativ sein, 2) auch die Kontrollstichprobe muß repräsentativ sein, und 3) die A-Merkmalsstichprobe und Kontrollstichprobe müssen vergleichbar sein. In allen drei Bereichen gibt es Quellen für Fehlermöglichkeiten, wofür wir einige Beispiele geben.

Fehlerquellen der Repräsentativität :

Man macht eine Telefonblitzumfrage zu einem politischen Thema. Damit erreicht man jedoch nicht die Nicht-Telefonbesitzer, also die ganz Armen; die Stichprobe ist

(vermutlich) nicht repräsentativ.

Ein Psychologe möchte ein Experiment z.B. zu Partnerschaftsgestaltung machen. Er inseriert in der Zeitung; Versuchspersonen melden sich. Welcher Typus von Person meldet sich auf ein solches Inserat? Wenn die Bezahlung der Versuchsperson nicht sehr gut ist (was sie meistens nicht ist), wird die Stichprobe nur die besonders an geistigen Auseinandersetzungen Motivierten und an Wissenschaft Interessierten enthalten. Die Stichprobe ist (vermutlich) nicht Repräsentativ. - Noch schlimmer: der Psychologe führt seine Untersuchung nur an Studenten, gar nur an Studenten seines Faches durch, die sich freiwillig melden. Und schließt daraus Zusammenhänge, die für alle Personen der Population gelten sollen.

Eine Soziologin schickt zur Erhebung der Arbeitsbelastung von Hausfrauen Fragebögen aus. 20% werden zurückgeschickt; welche Hausfrauen füllen ihn aus und schicken ihn zurück? Womöglich die bildungsmäßig gehobeneren, engagierteren, zugleich auch die, die weniger von Arbeit und Hausfrauenrolle belastet sind. Die Stichprobe ist (vermutlich) nicht repräsentativ.

Fehlende Kontrollgruppe: Ein typischer Fehler in populären nichtwissenschaftlichen Erfahrungsauswertungen ist, daß ein Vergleich mit einer Kontrollgruppe überhaupt fehlt (vgl. die obigen Beispiele zur Irrelevanz). Z.B. ermangeln die meisten Berichte über die Erfolge von Naturheilpraktiken oder sonstigen esoterischen Heilpraktiken einer gezielten Überprüfung durch Kontrollgruppen: wenn bei einer oder auch 10 Personen nach Anwendung der 'Pendelmethode' (Krankheiten wegpendeln) tatsächlich die Krankheit später spürbar nachließ, so sagt dies wenig aus, da man überhaupt nicht weiß, ob das Nachlassen der Krankheit nicht auch ohne das Heilungspendeln, durch normale Gesundheit, gekommen wäre, ob nicht ein Placebo-Effekt wirksam war (der starke Glauben an die Heilungskraft kann u.U. selbst die Gesundheit beschleunigen), daß der 'Heilpendler' nicht der kranken Person zugleich andere Dinge, z.B. gesünder leben, verordnete, die er in seinem Bericht verschwieg ('versteckte Variablen'), die aber in Wirklichkeit die wirksamen Faktoren waren, usw. All dies könnte mit einem Mal ausgeschlossen werden, wenn ein Kontrollgruppenvergleich angestellt würde (d.h. statistisch geprüft würde, ob die Genesungsrate bei vergleichbaren Personen mit gleicher Krankheit ohne 'Pendeln' geringer oder etwa gleich ist). Abgesehen davon werden vielen populären Darstellungen nur über die positiven Fälle berichtet, die negativen werden nicht erwähnt (ein noch 'primitiverer' Fehler). Nur eine kontrollierte statistische Untersuchung an vielen Personen, mit Merkmals- und Kontrollgruppe könnte einigermaßen verlässlichen Aufschluß über die Wirksamkeit geben; aber leider wehren sich die meisten 'Esoteriker' dagegen, auf diese Weise untersucht zu werden.

Fehlende Vergleichbarkeit mit Kontrollgruppe: Die Einschleusung von versteckten Faktoren in wissenschaftlichen Experimenten kann die Vergleichbarkeit

von Merkmalsgruppe und Kontrollgruppe erheblich verzerren. Ein Beispiel: eine neue Unterrichtsmethode soll mit einer herkömmlichen verglichen werden - getestet durch einen Leistungsvergleich zweier Schülergruppen (Merkmalsgruppe: neue Unterrichtsmethode; Kontrollgruppe: herkömmliche Unterrichtsmethode). Doch waren die Lehrer in beiden Gruppen auch gleich gut? Und insbesondere: waren die Lehrer der neuen Unterrichtsmethode vielleicht viel eher motiviert als die nach der herkömmlichen Methode unterrichtenden? Wußten die Schüler vom Experiment, waren sie -gegeben dieses Wissen - dadurch unterschiedlich motiviert? All dies kann zu erheblichen Verzerrungen führen, die falsche Signifikanzhypothesen zur Folge haben.

Interaktion Untersuchungsleiter-Versuchsperson: Ein weiterer Fehlerquellentypus, der auch Nichtrepräsentativität bewirkt, liegt in gewissen durch die Interview- oder Experimentsituation eingeschleusten versteckten Faktoren. Z.B. kann ein Befragungsergebnis stark von Person und Verhalten des Interviewers abhängen: so ergab eine Umfrage, wie D. Hull berichtet, zum Thema der politischen Rechte von Negern sehr unterschiedliche Ergebnisse, abhängig davon, ob ein Neger oder ein Weißer der Interviewer war. Bekannt ist auch der sogenannte *Rosenthal-Effekt*, wonach die Beurteilung des Ergebnisses eines Experiments, z.B. zum Lernerfolg, von den Vorerwartungen des Versuchsleiters bzw. Auswerters abhängt. Rosenthal ließ Studenten (als Versuchsleiter) eine Reihe von Versuchspersonen (z.B. Schüler) bzgl. ihrer Leistung in einem Test als erfolgreich oder nichterfolgreich beurteilen, wobei einem Teil der Studenten vorher gesagt wurde, es handle sich um besonders leistungsstarke Vpn, einem anderen Teil der Studenten, es handle sich um leistungsschwache Vpn - tatsächlich waren die Vpn, wie in einem Vortest sichergestellt, alle gleich leistungsstark. Die Versuchsleiter mit der Vorinformation "leistungsstark" beurteilten die Vpn signifikant positiver hinsichtlich ihres Leistungserfolges als die Versuchsleiter mit Vorinformation "leistungsschwach".

3.3 Das Problem der Kausalität: Korrelation und Kausalität

Gegeben eine Gesetzhypothese ist vermutlich sowohl wahr wie relevant. D.h. es liegt, statistisch gesprochen, eine signifikante Korrelation vor. Sie ist sicher prognostisch relevant, kann zur Vorhersage verwendet werden. dennoch kann man daraus nicht unmittelbar schließen, daß auch ein Kausalzusammenhang besteht. Und zwar aus zwei Gründen.

3.3.1 Versteckte Variablen

Angenommen es besteht eine stark positive Korrelation $Korr(A,B)$. Das ist noch

keine Rechtfertigung für die These, daß zwischen A und B ein Kausalzusammenhang. $\text{Korr}(A,B)$ kann eine sogenannte "Scheinkorrelation" sein - man sollte aber besser von *Scheinkausalität* sprechen; die Korrelation besteht ja wirklich: zwischen A und B besteht dann überhaupt keine kausale Beziehung, sie sind Wirkungen einer gemeinsamen versteckten Ursache C, und die Korrelation zwischen A und B kommt nur durch die Kausalbeziehung C-B und C-A zustande.



Einige Beispiele:

(1) Zwischen dem Fall des Barometers und dem Kommen des Sturms besteht hohe Korrelation.

Da das Kommen des Sturms auf den Barometerfall folgt, könnte man von der Richtung her meinen, letzterer wäre Ursache des ersteren. Aber beide sind Ursache einer dritten Variablen, dem Druckabfall in der Atmosphäre.

(2) Zwischen dem Grad der positiven Einstellung zur Firma und der psychischen Gesundheit von Arbeitern wurde eine hohe positive Korrelation gemessen.

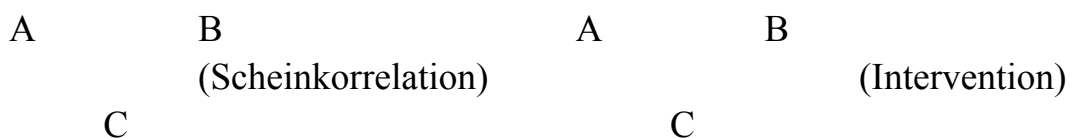
Heißt das, die dauernd über die Firma Nörgelnden sind die psychisch Labilen oder Defekten - die Firma ist unschuldig? Es stellte sich heraus, daß beide gemeinsame Wirkung der versteckten Ursache der Arbeitsplatzbelastung sind: die Arbeiter mit schlechten Arbeitsbedingungen klagten natürlich über die Firma, zugleich leidete ihre psychische Gesundheit aufgrund der fortwährende Belastung. - Ein ganz anderes Bild ergibt sich (das Beispiel stammt von Lazarsfeld).

Es gibt eine bekannte Methode, um Scheinkorrelationen statistisch aufzudecken. Angenommen, es gilt $\text{Korr}(A,B)=\text{pos.}$ Wenn nun C eine gemeinsame Ursache ist, dann muß die Korrelation zwischen A und B bei festgehaltenem C - die schon erwähnte Partialkorrelation $\text{Korr}(A,B/C)$ - verschwinden. D.h. in unserem Beispiel (2): wenn man nur die Arbeiter unter guten Arbeitsbedingungen untersucht, verschwindet die Korrelation zwischen Einstellung zur Firma und psychischer Gesundheit; ebenso, wenn man nur Arbeiter unter schlechten Arbeitsbedingungen untersucht. Allgemeiner: wenn man schon weiß, daß C vorliegt, oder daß nicht-C vorliegt, wird die Wahrscheinlichkeit von A durch das Wissen um B nicht mehr weiter erhöht, da zwischen B und A eben keine kausale Beziehung existiert.

Diese Methode der Ausschließung von Scheinkorrelationen würde verlangen, daß man *alle möglicherweise relevanten Hintergrundparameter* untersucht und prüft, ob durch sie die gemessene Korrelation zwischen A und B zum Verschwinden gebracht wird. Man sieht, daß für die kausale Interpretation von statistischen Korrelationen ein *enormes Maß an theoretischen Hintergrundannahmen* nötig ist - nämlich eine Annahme über alle möglicherweise relevanten Parameter.

Ferner hat diese Methode eine Reihe weiterer Schwächen. Erstens versagt sie bei hohen oder annähernd 1-Wahrscheinlichkeiten. Zweitens kann sie im Viel-Variablen-Fall versagen (z.B. bei multipler Regression), wo es passieren kann, daß sich alle Variablen in ihren Korrelationen gegenseitig zum Verschwinden bringen. Drittens kann mit dieser Methode nicht zwischen Scheinkorrelation und Intervention unterschieden werden - was wir genauer ausführen wollen.

Bei der *Intervention* schiebt sich die dritte Variable C, die die Korrelation zwischen A und B zum Verschwinden bringt, als Mittlerursache zwischen A und B.



Die beiden Fälle sind statistisch völlig gleichartig. Im Fall der Intervention muß jedoch von einem indirekten Kausalzusammenhang zwischen A und B gesprochen werden. Welcher Fall vorliegt, kann wiederum nur durch zusätzliches Wissen über Kausalrichtungen und Kausalmechanismen entschieden werden, welches nicht in den statistischen Daten enthalten ist, sondern in den Hintergrundtheorien. Einige Beispiele für Interventionen:

(3) Zwischen dem Familienstand berufstätiger Frauen (A=verheiratet) und der Häufigkeit ihres Fernbleibens von der Arbeit (B) maß man hohe Korrelation. Sind die verheirateten Frauen so faul? Keineswegs: als intervenierende Variable bemerkte man das Ausmaß zusätzlicher Belastung durch Hausarbeit (C). A bewirkt hohes Ausmaß an C = zusätzliche Hausarbeit, und deshalb bleiben die verheirateten Frauen häufiger im Betrieb von der Arbeit fern (hohes B). Sie sind die fleißigeren, nicht die faulen. (Beispiel von H. Zeisel).

(4) Ein interessantes Beispiel stammt von Nancy Cartwright. In einer amerikanischen Universität stellte man fest, daß bei weiblichen Studenten (=W) die Aufnahmequote (=A) an die Universität (es gibt hier Aufnahmeprüfungen) niedriger ist als bei männlichen. Man kann hier gut die möglichen Kausalvermutungen und ihre

Widerlegung durch die Untersuchung von Hintergrundvariablen erkennen.

1. Vermutung: Frauen sind einfach dümmer, fallen deshalb bei den Aufnahmeprüfungen häufiger durch. Hält man die Variable des IQs daher konstant, so müßte die Korrelation verschwinden, d.h. $\text{Korr}(A, W/IQ)=0$ gelten. (D.h., wenn man gleichintelligente Frauen mit Männern vergleicht, dürfte zwischen beiden in der Aufnahmequote kein Unterschied mehr bestehen). Ergebnis:

$\text{Korr}(A, W/IQ)=$ weiterhin positiv; das kann also nicht die gemeinsame Ursache sein.

2. Vermutung: Das Universitätssystem bevorzugt einfach Männer (Patriarchatshypothese). Gemeinsame Ursache wäre jetzt also $P =$ patriarchalische Gesellschaftsstruktur. Hier hat man ein typisches Beispiel, wo man nicht mehr weiteruntersuchen kann, weil man die Variable P eben nicht variieren kann (man kann mit Gesellschaften nicht 'kontrolliert experimentieren').

Dennoch zeigt eine Information, daß auch diese zweite Vermutung nicht Ursache dieser Korrelation ist, denn:

3. Man beobachtet, daß Frauen zur Inskription in den beliebteren Fächern tendieren als Männern, d.h. in jenen Fächern, die (wegen der hohen Nachfrage) niedrigere Aufnahmequoten haben. Als man die Hintergrundvariable des Studienfaches konstant hielt ($=F$), verschwand die Korrelation - d.h. in jedem einzelnen Fach ergab sich keinerlei Unterschied zwischen männlich- und weiblich-Aufnahmequoten; die Korrelation resultierte nur daraus, daß Frauen gewisse Fächer (die 'beliebteren' wie Psychologie, Germanistik usw.) häufiger studierten als Männer. Dabei ist (F) hier intervenierende Variable: Frauen tendieren häufiger zur Inskription von eher überbelegten Fächern, was Ursache für die niedrigeren Aufnahmequoten ist.

(5) Ein anderes Beispiel, wo es wichtig wird zu wissen, ob Scheinkorrelation oder Intervention vorliegt: zwischen Kaffeetrinken ($=K$) und Herzleiden ($=R$) entdeckte man eine Korrelation, die verschwand, als man Rauchen als dritte Variable einführte. Ist nun Kaffeetrinken Ursache des Rauchens (verleitet es zum Rauchen), und Rauchen direkte Ursache des Herzleidens? - dann läge Intervention vor. Oder ist umgekehrt Rauchen Ursache des Kaffeetrinkens (verleitet es zum Kaffeetrinken) und zugleich Herzleidens? - dann läge Scheinkorrelation vor.

Abschließend sei erwähnt, daß es - über Scheinkorrelation und Intervention hinaus - auch den Fall der *statistischen Scheinunabhängigkeit* gibt, wo man zunächst $\text{Korr}(A, B)=0$ mißt, bei Einführung einer Hintergrundvariablen C dann aber sowohl im Fall C wie im Falle $\neg C$ eine positive Korrelation zwischen A und B mißt, d.h. $\text{Korr}(A, B/C)=\text{pos.}$ und $\text{Korr}(A, B/\neg C)=\text{pos.}$ (Das sogenannte "Simpson-Paradox"). Wir gehen darauf nicht mehr ein.

3.3.2 Kausalrichtung

Selbst in dem hypothetischen (meistens unrealistischen) Fall, wo man weiß, daß außer A und B keine weiteren ("versteckten") Variablen mehr im Spiel sind, kann man aus dem Vorliegen einer hohen Korrelation $Korr(A,B)$ nicht direkt auf Kausalität schließen, weil durch die Korrelation nicht die Kausalrichtung festgelegt ist - also es ist nicht festgelegt, was als Ursache und was als Wirkung anzusehen ist.

Korrelationen sind immer symmetrisch (das geht aus ihrer mathematischen Definition hervor). Wenn also zwischen zwei Merkmalen A, B eine Korrelation $Korr(A,B)$ besteht, so ist $Korr(A,B)$ positiv g.d.w. $Korr(B,A)$ positiv ist (es gilt sogar: $Korr(A,B)=Korr(B,A)$). Angenommen es besteht überhaupt ein Ursache-Wirkungs-Zusammenhang zwischen A und B - dann gibt es im Falle $Korr(A,B)=\text{positiv}$ zwei einfache Möglichkeiten: A ist ein Kausalfaktor für B, oder B ein Kausalfaktor für A. Schließlich gibt es die komplexe Möglichkeit, daß A und B in Form einer *Rückkopplung* kausal aufeinander einwirken, wobei das Ergebnis dann ganz von der Art der Rückkopplung abhängt. Einige Beispiele (aus tatsächlichen Untersuchungen oder Berichten stammend):

(6) Höhe des IQ und Höhe des sozialen Status sind positiv korreliert.

Entscheidende Frage: was ist hier Ursache, was Wirkung? Ist IQ Produkt der Erziehung und Bildung, welche vom Status abhängt? Oder denkt der Verfechter der Korrelation an eine genetische Elite?

(7) Aggressive Personen sehen gerne aggressive Filme.

Die wichtige Frage ist aber: Werden Personen durch aggressive Filme aggressiv? Oder umgekehrt: Sehen aus anderen Gründen aggressive Personen gerne aggressive Filme, wobei sie sich dabei abregieren?

(8) Interesse für Computer korreliert mit Desinteresse an sozialen Beziehungen.

Die wichtige Frage ist hier: Versuchen die sozial Beziehungsunfähigen ihren Mangel durch intensive Beschäftigung mit technischen Dingen wie Computer auszugleichen? (Psychologische Deutung). Oder ist die Sache einfach so, daß die meisten Computerbeschäftigten durch ihre Arbeit so erfüllt werden, daß sie aus natürlichen Gründen ihr Engagement in sozialen Beziehungen reduzieren? (Rationale Deutung). Oder aber, verändert die Beschäftigung mit Computer tatsächlich mit der Zeit auf den Charakter?

Man ist in obigen Beispielen schnell geneigt, eine Korrelation in der Richtung zu lesen, die der *eigenen Auffassung* entspricht. Wichtig ist: in welche Richtung die Kausalbeziehung geht, kann nicht durch statistische Korrelationen erhoben werden, da diese immer symmetrisch sind. Echte wissenschaftliche Haltung kann nur in Zurückhaltung bzgl. der Kausalrichtung bestehen, sofern man nicht zusätzliche Gründe für die Kausalrichtung anführt. Solche Gründe sind *theoretischer Natur*, also resultieren aus Hintergrundtheorien, z.B. naturwissenschaftliches Hintergrundwissen über Kausalmechanismen. Im Fall von Beziehungen zwischen zeitlich getrennten Momentanereignissen kann nur das frühere die Ursache des späteren sein, denn *Kausalbeziehungen sind zeitlich vorwärtsgerichtet*. Schwieriger wird es schon bei zeitlich nicht lokalisierten Dispositionsmerkmalen. Man muß hier herausfinden, was die unabhängige und was die abhängige Variable ist. Auch da gibt es einfache Fälle. Wenn wir z.B. Korrelationen wie

- (9) Geschlecht korreliert mit gewissen geschlechtspezifischen Merkmalen
- (10) Luftverschmutzung korreliert mit Häufigkeit von Atemwegserkrankungen

vor uns haben, ist aufgrund naturwissenschaftlicher Zusammenhänge die Richtungsfrage klar. Im Fall (9) ist klar, daß wenn überhaupt ein Kausalzusammenhang vorliegt, dann das Geschlecht die Ursache sein muß, weil es biologisch vorgegeben ist; und im Fall (10) ist aufgrund naturwissenschaftlichem Hintergrundwissen der Kausalmechanismus bekannt, der von Luftverschmutzung zu Atemwegserkrankung führt. - Jedoch in den obigen Beispielen (6), (7), (8) ist die Frage der Kausalrichtung sehr fraglich, sofern überhaupt eine Kausalbeziehung besteht. Ein Beispiel für eine folgenschweren Fehlinterpretation einer Korrelation in der Kausalrichtung ist der historisch bekannte Fall des Bankiers John Law, der aus der Korrelation zwischen der Menge des im Umlauf befindlichen Papiergeldes und dem Reichtum der Nation schloß, man könne durch In-Umlauf-Setzen von ungedecktem Papiergeld den Reichtum der Nation erhöhen. Der Zusammenbruch seiner Banken, darunter die Nationalbank, war die Folge.

Abschließend muß gesagt werden, daß heutzutage in den Medien statistische Korrelationsergebnisse zumeist unkritisch sofort als Kausalergebnisse gedeutet werden - hier wäre es Aufgabe der Wissenschaftstheorie, solche Fehlschlüsse kritisch zu korrigieren bzw. auf die Problematik überhaupt erst aufmerksam zu machen.

**3.4 Erklärung, Begründung und Voraussage: Der deduktiv-deterministische und der induktiv-statistische Fall*

Zunächst ist zu bemerken, daß der Erklärungs-begriff im Alltag mehrdeutig verwendet wird: erklären, warum etwas der Fall war, erklären, wozu etwas der Fall ist oder sein soll, die Bedeutung eines Wortes/Satzes erklären, usw. Der in den Wissenschaften primär relevante Erklärungstyp, auf den wir uns hier beschränken, ist die Warum-Erklärung.

Wissenschaftliche Erklärungen, Begründungen und Voraussagen von Einzelereignissen bzw. singulären Sätzen besitzen ganz allgemein immer folgende Bestandteile:

(a) ein oder mehrere Gesetze, die deterministischer oder statistischer Natur sein können - zusammengefaßt in (der Satzmenge) G

(b) ein oder mehrere singuläre Sätze, die zusammengefaßt das Antecedens A der Erklärung/Voraussage/Begründung genannt werden. (A umfaßt singuläre Anfangs- und Randbedingungen).

- G und A zusammen werden das Explanans (Demonstrans, Prognosens) genannt; und

(c) einem singulären Satz, dem Explanandum (Demonstrandum, Prognosendum) E, welches die zu erklärende (begründende, vorauszusagende) Tatsache ausdrückt.

Erklärungen (Begründungen, Voraussagen) mit deterministischen Gesetzen sind deduktive Argumente. Hier ist das Explanandum E eine logische Folge des Explanans. Die Theorie dieser Art von Erklärungen geht auf Carl Gustav Hempel zurück, er sprach von deduktiv-nomologischen Erklärungen, ein Ausdruck, der sich eingebürgert hat. Dennoch ist es terminologisch besser, von deduktiv-deterministischen Erklärungen, kurz DD-Erklärungen zu sprechen, weil die Gesetze bei solchen Erklärungen ("nomos"- "Gesetz") immer deterministischer Natur sind. Erklärungen (Begründungen, Voraussagen) mit statistischen Gesetzen sind dagegen induktive Argumente bzw. 'Quasiargumente' (also keine echten Argumente, bei denen es eine 'abspaltbare' Konklusion gibt - s.u.). Das Explanandum ist hier nicht logische Folge des Explanandum, sondern ist lediglich aufgrund des Explanans mit einer gewissen induktiven Erwartungswahrscheinlichkeit erwartbar. Auch die Theorie dieser Erklärungstyps geht auf C.G.Hempel zurück - wir sprechen hier, mit ihm übereinstimmend, von induktiv-statistischen Erklärungen, kurz IS-Erklärungen. DD-Erklärungen und IS-Erklärungen haben, zusammengefaßt, unterschiedliche logische Struktur.

Wir wollen Explanans (Demonstrans, Prognosens) auch allgemein Prämissenmenge, und Explanandum (Prognosendum, Demonstrandum) allgemein Konklusion nennen - bei DD- wie IS- Erklärungen/Begründungen/Voraussagen. Dann kann man den Unterschied zwischen Erklärung, Begründung und Voraussage

ganz allgemein so definieren:

Ein DD- oder SI-Argument ist eine Voraussage oder Retrodiktion, wenn seine Prämissen zuvor bekannt waren und die Konklusion erst hinterher daraus abgeleitet wurde. Liegt dabei das Konklusionsereignis in der Zukunft, so liegt eine Voraussage vor, liegt es in der Vergangenheit, so liegt eine Retrodiktion vor. Voraussagen bzw. Retrodiktionen nennt man zusammenfassend auch ex-ante Begründungen. Eine ex-post Begründung oder auch "Erklärung im weiten Sinn" liegt dagegen vor, wenn das Konklusionsereignis zuvor schon bekannt war (beobachtet wurde) und seine Prämissen erst nachträglich hinzugefunden wurden. Eine ex-post-Begründung kann man auch Erklärung im weiten Sinn nennen. Eine Erklärung im engen Sinn (eigentlichen Sinn) liegt jedoch nur dann vor, wenn folgende Zusatzbedingung erfüllt ist: zwischen Antecedens und Explanandum besteht eine Kausalbeziehung, d.h. das Antecedens ist Ursache, Explanandum ist Wirkung. - Wir sehen also: Ex post-Begründungen (Erklärungen i.w.S.) unterscheiden sich von ex ante Begründungen (Voraussagen und Retrodiktionen) nur durch die pragmatischen Zeitumstände des Bekanntwerdens von Prämissen und Konklusion. Dagegen ist der Unterschied zwischen Begründungen und echten Erklärungen ein nicht pragmatischer, sondern substanzieller: Echte Erklärungen müssen eine Kausalbeziehung zwischen A und E enthalten. Man kann den Unterschied zwischen Begründungen und Erklärungen im engen Sinn auch so definieren: eine Sachverhalt zu begründen heißt, gute Glaubensgründe (Beweisgründe) für seine Wahrheit zu liefern; ein Argument zu liefern, aufgrund dessen die Wahrheit des Demonstrandums erhärtet oder bewiesen wird. Einen Sachverhalt zu erklären heißt aber mehr: Seinsgründe bzw. Ursachen für sein tatsächliches Bestehen bzw. Auftreten zu liefern, jene Faktoren zu nennen, die für das Eintreten des Sachverhalts kausal verantwortlich waren. (wobei man beim Erklären schon voraussetzt, daß das Erklärte tatsächlich der Fall ist). Diese Unterscheidung entspricht dem intuitiven Gebrauch von "Erklärung" und "Begründung". Z.B. laute die Frage: "warum ist das Wetter heute am Ort x schlecht?" Die Antwort "weil der Wetterbericht gestern für Ort x Schlechtwetter voraussagte" ist sicher akzeptabel als Begründung des Satzes (vorausgesetzt die Verlässlichkeit des Wetterberichtes). Aber es erklärt doch nicht, warum das Wetter dort schlecht ist - erklären würde eine Antwort wie "weil am Ort x gestern aufgrund einer klimatischen Nordströmung Tiefdruck herrschte...(usw.)."

Wir behandeln nun den deterministisch-deduktiven Fall genauer. Eine einfache DD-Begründung sieht so aus:

G: $\forall x(Fx \rightarrow Gx)$	Alle Metalle (=Fx) leiten Strom (=Gx)
A: <u>Fa</u>	<u>Dieses Ding (=a) ist aus Metall</u>

E: Ga (Daher:) Dieses Ding leitet Strom

Wenn wir in obigem Beispiel etwa eine Person sehen, die mit einem vollmetallischen Schraubenschlüssel ein an die Steckdose angeschlossenes Radio reparieren will, und wir warnen "Achtung - der Schraubenzieher leitet den Strom", so hätten wir das Argument als Voraussage verwendet. Wenn die Person sich jedoch bereits elektrisiert hat, und wir ihr dies (oder ihren Anverwandten) damit erklären, daß der Schraubenschlüssel aus Metall war, liegt eine ex-post-Begründung vor, die hier auch eine echte Erklärung ist, weil das Metallisch-sein eine Ursache des Stromleitens ist.

Aber nicht jede ex post facto Begründung ist eine echte Erklärung. Zwei Beispiele:

G: Immer, wenn der Barometerstand fällt, wird ein Sturm kommen

A: Der Barometerstand fiel.

E: Daher wird ein Sturm kommen

oder:

G: Jeder, der progressive Paralyse hat, wurde zuvor von Syphilis infiziert.

A: Peter hat progressive Paralyse.

E: Daher: Peter wurde von Syphilis infiziert.

Im ersten Beispiel ist A bloß ein Indikator, aber nicht die Ursache von E - A und E sind beide korrelierte Wirkungen einer dritten, gemeinsamen Ursache - der Druckabfalles in der Atmosphäre. Im zweiten Beispiel ist A nicht Ursache von E, sondern eine Wirkung von E - denn A liegt zeitlich nach E, wobei zusätzlich E die einzig mögliche Ursache für A ist, weshalb von der Wirkung A zeitlich zurück auf die Ursache E geschlossen werden kann. Die beiden Fälle - (1.) Verletzung der (zeitlichen) Kausalrichtung und (2.) versteckte gemeinsame Ursache - sind zugleich schon die wichtigsten Gründe für die Tatsache, daß nicht jede Voraussage eine potentielle Erklärung ist. Es gibt auch Gründe, warum umgekehrt nicht jede Erklärung eine potentielle Voraussage ist (z.B. im Zusammenhang mit dem derzeit aktuellen 'deterministischen Chaos'), worauf hier nicht eingehen können. Jedenfalls ist nach heutigem Erkenntnisstand Hempels ursprüngliche These der strukturellen Gleichartigkeit von Erklärung und Voraussage nicht haltbar.

DD-Erklärungen kommen hauptsächlich in den Naturwissenschaften, aber auch z.B. in den Wirtschaftswissenschaften vor:

G: Für alle freie Pendelsysteme mit Reibungskraft gilt: Federkraft + Reibungskraft + Trägheitskraft = Null

A: a ist ein freies Pendelsystem mit Reibungskraft; und zur Zeit 0 hatte es Anfangsauslenkung q und Anfangsgeschwindigkeit p .

E: Daher hat beschreibt die Auslenkung von a für alle Zeiten genau die und die Bewegungskurve (was durch Lösung der in G stehenden Differentialgleichung und Einsetzung der Anfangsbedingungen unter Benutzung der mathematischen Analysis logisch gefolgert werden kann).

G: Für alle freie Tauschmärkte x: der Gleichgewichtspreis eines Gutes y von x ist umso höher, je geringer seine zum Tausch angebotene Gütermenge (das Angebots-Nachfrage-Gesetz der klassischen Tauschwirtschaft)

A: Am freien Tauschmarkt a verringerte sich zur Zeit t durch Ernteeinbußen das Getreideangebot drastisch.

E: Daher erhöhten sich bald nach t (nach einer 'Gleichgewichtseinpindelungszeit') am Tauschmarkt a die Preise von Getreide.

In den Human-, Sozial- und Geschichtswissenschaften lassen sich deterministische Gesetze - wegen der Komplexität der Faktoren - kaum ausfindig machen. Daher spielen hier IS-Erklärungen eine wesentliche Rolle. Gemäß Hempel haben Erklärungen folgende Form:

G: $p(G/F)=95\%$	95% aller mit Plasmodium vivax infizierten Personen erkrankten an Malaria
A: Fa	Peter wurde mit P.v.infiziert
===== [95%]	===== [95%]
E: Ga	(Daher:) Peter erkrankte an Malaria.

Der in [] stehende Wert ist dabei die induktive Erwartungswahrscheinlichkeit, welche sich auf den Einzelfall a bezieht, und somit keine statistische Häufigkeit darstellt, sondern unseren Glaubensgrad bzgl. des Einzelereignisses Ga angibt (also eine epistemische Wahrscheinlichkeit. In einem solchen IS-Schema wird also die statistischen Häufigkeit als epistemischer Erwartungswert auf einen einzelnen Fall übertragen. Das Analogon der deduktiven Folgerungsbedingung im DD-Fall ist im IS-Fall dabei die Bedingung, daß die Wahrscheinlichkeit hoch sein muß. Andernfalls wird es nicht vernünftig sein, ein Gesetz G für eine IS-Begründung (Voraussage oder Erklärung) eines Einzelereignisses einzusetzen.

Eine in der Wissenschaftstheorie vielfach diskutierte Frage ist es, wie hoch der Wahrscheinlichkeitswert sein muß, damit man gerade noch von einer Erklärung/Voraussage/Begründung sprechen kann. Von einer Begründung oder Voraussage wird man verlangen, daß die Wahrscheinlichkeit $p(E/A)$ mindestens

größer als 0,5 bzw. $p(\neg E/A)$ ist. Was Erklärungen betrifft, so meinen manche Autoren, daß man auch noch dann von Erklärung sprechen könne, wenn $p(E/A)$ bloß größer ist als $p(E)$, aber nicht notwendigerweise größer als 0,5. Wesley Salmon (1984) hat ziemlich überzeugend argumentiert, daß im Fall von Kausalerklärungen die wichtige Frage eigentlich gar nicht in der Höhe der Wahrscheinlichkeit liegt, sondern in der Forderung, A müsse alle statistisch relevanten Kausalfaktoren enthalten. Wir können auf diese Fragen hier nicht näher eingehen; vgl. hierzu Schurz (Hg., 1988).

Ferner ist zu erwähnen, daß Hempels ursprüngliche Modellen eine Reihe von nötigen Zusatzbedingungen an rational akzeptierbaren Erklärungen nicht enthielt, die wir jedoch völlig unproblematisch mitaufnehmen können. Z.B. wissen wir, daß ein (deterministisches oder statistisches) Gesetz nur akzeptierbar ist, wenn es die Relevanzbedingung enthält - welche in Hempels Modellen völlig fehlte, worauf insbesondere dann Salmon (1971) hinwies, der sein 'alternatives' Modell das 'Statistische Relevanzmodell' nannte. Auf diese sowie weitere Bedingungen, die z.B. deduktive Relevanzforderungen betreffen (vgl. Schurz 1983, 1988), können wir hier nicht weiter eingehen.

IS-Begründungen (bzw. Erklärungen) unterscheiden sich von DD-Begründungen (Erklärungen) jedoch durch eine zusätzliche Bedingung, die für ihr Verständnis ganz zentral ist. Allgemein gilt:

Eine DD-Erklärung "G,A/E" ist rational akzeptierbar, wenn G ein als rational akzeptierbares deterministisches Gesetz ist, A und E rational akzeptierbare singuläre Sätze sind, und E aus G und A (in deduktiv relevanter Weise) logisch folgt.

Eine IS-Erklärung dagegen ist nicht schon dann als gültig akzeptierbar, wenn G ein rational akzeptierbares statistisches Gesetz ist, A und E rational akzeptierbare singuläre Sätze sind, und E durch G und A hochwahrscheinlich gemacht wird. Zusätzlich muß gelten: A enthält die maximale für E statistisch relevante Information über das erklärte Individuum a relativ zum gegebenen Hintergrundwissen W. Hempel nannte dies die Forderung der "maximalen Bestimmtheit".

Die Notwendigkeit der maximalen Bestimmtheitsforderung im IS-Fall läßt sich an unserem obigen Beispiel leicht einsehen. Es gibt nämlich eine Erbeigenschaft H, bei deren Vorliegen Personen gegen Plasmodium vivax weitgehend immun werden und nur in 1% der Fälle an Malaria erkranken. Es gilt also auch das Gesetz: $p(G/F \wedge H) = 1\%$. Angenommen, wir wissen über Peter zusätzlich, daß er diese seltene Erbeigenschaft hat (H_a). Dann kommen wir mit dem erweiterten Gesetz zu einer Begründung, die E nun plötzlich statt 95%-wahrscheinlich nur 1%-unwahrscheinlich macht.

$p(G/F \wedge H) = 1\%$	"
$F_a \wedge H_a$	"

===== [1%] ===== [99%]
 Ga ¬Ga

Wir würden zu widersprüchlichen Wahrscheinlichkeitsbewertungen, ja sogar (wie rechts ersichtlich) zu widersprüchlichen Prognosen gelangen, wenn wir beide Begründungen/Erklärungen zuließen. Tatsächlich aber dürfen wir nur eine davon zulassen, nämlich jene, die die gesamte für Peters Malariaerkrankung statistisch relevante Information über Peter enthält. Das ist in diesem Fall die Erklärung mit dem Antecedens $F_a \wedge H_a$. Aus diesem Grund ist die rationale Akzeptierbarkeit einer IS-Erklärung/-Begründung immer relativ zum gerade gegebenen Hintergrundwissen; kommt neues Wissen hinzu, so kann aus einer akzeptierbaren eine nicht länger akzeptierbare IS-Erklärung werden, auch wenn alle ihre Prämissen wahr sind. Dies ist ein entscheidender Unterschied zu DD-Erklärungen und überhaupt allen logischen Folgerungsrelationen. Eine logische Folgerungsrelation "aus A folgt B" gehorcht nämlich dem Prinzip der Prämissenverstärkung (damit zusammenhängend: 'Abspaltbarkeit' der Konklusion): "wenn A wahr ist, ist mit Sicherheit B wahr - egal was sonst noch wahr ist". Genau dies gilt bei induktiven Relationen nicht - es kann sein daß A B wahrscheinlich macht, A mit C zusammen aber B unwahrscheinlich machen. Daher die essentielle Hintergrundrelativität und die maximale Bestimmtheitsforderung bei IS-Erklärungen/-begründungen. Wir erkennen daraus eine weitere, in der Auswertung statistischen Wissens oft mißachtete Bedingung: die Anwendung eines statistischen Gesetzes auf einen Einzelfall ist nur dann erlaubt, wenn das Gesetzesantecedens tatsächlich nicht nur irgendeine, sondern die gesamte für das Explanandum statistisch relevante Information über diesen Einzelfall enthält. Wenn wir z.B. herausfinden, daß eine Unterrichtsmethode in 90% aller Fälle Vorteile bringt, und wir dieses Wissen dann auf den Schüler Peter anwenden, so müssen wir genau prüfen, ob es nicht spezifischere Informationen gibt, die den Wahrscheinlichkeitswert dieses statistischen Gesetzes unterlaufen. Es könnte z.B. sein, daß die Unterrichtsmethode nur gewisse Lerntypen (z.B. visuelle) Vorteile bringt, zu denen insgesamt 95% aller Schüler zählen ausmachen, daß aber Peter zufällig ein Lerntyp ist, der in die restlichen 10% fällt (z.B. ein auditiver), wo diese Unterrichtsmethode Nachteile bringt. Die Anwendung des Gesetzes auf Peter wäre dann nicht nur illegitim; sie würde ihm praktisch Schaden bringen.

Abschließend noch zum Zusammenhang Begründung und Bestätigung. Im DD-Fall ist der Zusammenhang offenbar völlig symmetrisch. Nehmen wir ein deterministisches Gesetz $\forall x(Fx \rightarrow Gx)$, das zusätzlich relevant sein soll. Dann wird durch jede es bestätigende Instanz $F_a \wedge G_a$ zugleich eine DD-Erklärung " $\forall x(Fx \rightarrow Gx)$, F_a/G_a " möglich. D.h. hier gilt uneingeschränkt: ein deterministisches Gesetz wird umso mehr bestätigt, je mehr erfolgreiche Begründungen (Voraussagen, Erklärungen) es leistet. Dagegen wird es falsifiziert, wenn es nur eine Begründung eines Satzes E

liefert, der das Gegenteil dessen ist, was beobachtet wurde.

Im IS-Fall sind Erklärung und Bestätigung dagegen nicht unmittelbar gekoppelt, denn Bestätigung von statistischen Gesetzen ist nur mittels Stichproben möglich, während sich unser IS-Erklärungsmodell auf Einzelereignisse bezog. Daher habe ich andernorts (z.B. 1988) ein Modell der IS-Stichprobenerklärung (-begründung) entwickelt - ein Beispiel:

G: Die Wahrscheinlichkeit, daß das Kind blaue Augen hat, wenn beide Elternteile jeweils ein Braun- und ein Blau-Gen haben, ist $1/4$. (Das Mendelsche Gesetz)

A: Diese Stichprobe von 100 Kindern haben Eltern, bei denen beide Elternteile jeweils ein Blau- und ein Braun-Gen haben.

=====_[95,5%]

E: Zwischen 16,5 und 33,5 Kinder davon haben blaue Augen.

Bezogen auf solche Stichprobenerklärungen gilt ebenso: ein statistisches Gesetz wird umso besser durch Stichproben bestätigt, je mehr erfolgreiche Stichprobenbegründungen es leistet. Dagegen wird es geschwächt (falsifiziert im Wahrscheinlichkeitssinn), wenn es nur eine Stichprobenbegründung liefert, bei der die tatsächlich prognostizierte Stichprobenhäufigkeit außerhalb des prognostizierten Konfidenzintervalls liegt.

3.4.1. Erklärung menschlicher Handlungen. Begründungen in der Mathematik und in präskriptiven Disziplinen.

Können mit diesen beiden - den empirischen Wissenschaften angepaßten - Erklärungsmodellen auch menschliche Handlungen erklärt werden? Dies war eine wissenschaftstheoretisch vieldiskutierte Frage - bekanntlich vertraten Philosophen wie Wilhelm Dilthey u.a. die Ansicht, im Bereich der Geisteswissenschaften ginge es nicht um Erklären aufgrund Gesetze, sondern um Verstehen individueller Situationen. Heute stehen die Begriffe des Erklärens und Verstehens keinesfalls mehr in dem Gegensatz, in dem man sie im 19. Jahrhundert sah (vgl. Schurz 1988). - Zunächst sei bestritten ist, daß menschliche Handlungen sicher nicht deterministisch, sondern bestenfalls statistisch erklärbar sind. Ein von mehreren Autoren in unterschiedlichen Nuancen vorgeschlagenes Modell ist das folgende Modell der deskriptiven rationalen Erklärung von Handlungen:

G:

Für alle Personen x gilt: wenn

(a) wenn x im Zeitintervall t Ziel Z erreichen will

(b) x in t bzgl. Z einen homogenen Willenszustand hat (i.e. nicht zugleich ein Ziel Z' will, das mit Z unverträglich ist)

(c) x in t glaubt, daß die Durchführung von Handlung H in t eine notwendige Bedingung für die Realisierung von Z ist

(d) x in t weder durch äußere Zwänge noch durch innere unbewußte Zwänge an der Realisierung von H innerhalb t gehindert wird,

dann wird x mit hoher statistischer Wahrscheinlichkeit innerhalb t die Handlung H ausführen.

A:

Die Salzburgerin Susi will zu Semesterbeginn Philosophie studieren, und sie hat diesbezüglich einen homogenen Willenszustand (will also nicht zugleich etwas damit Unvereinbares studieren) und sie glaubt, daß die Inskription des Philosophiestudiums an der Universität Salzburg zu Semesterbeginn eine notwendige Bedingung dafür ist, und wird weder durch äußere Zwänge (Finanzprobleme, Eltern, etc.) noch durch innere unbewußte Zwänge (z.B. gestörtes Verhältnis zum Lehrer im Philosophiegymnasialunterricht) daran gehindert, zu Semesterbeginn an der Salzburger Uni Philosophie zu inskribieren.

E:

Daher ist mit hoher Wahrscheinlichkeit zu erwarten, daß Susi an der Salzburger Universität zu Semesterbeginn Philosophie inskribieren wird.

Dieses Modell ähnelt dem Weingartnerschen echter teleologischer Erklärungen (die zusätzliche Klausel "homogener Willenszustand" ist eine wichtige, von mir eingeführte Zusatzbedingung). Es handelt sich hier um eine deskriptive Handlungserklärungen, weil erklärt wird, warum die Handlung tatsächlich eintrat, und nicht etwa, warum es vernünftig gewesen wäre, so zu handeln. Im letzteren Fall sprechen wir von normativer Handlungsbegründung - die Idee dazu geht auf W. Dray zurück. Obige Handlungserklärung heißt rational, weil sie sich auf rationale, d.h. von der Person willentlich geplante Handlungen bezieht - unter Ausschluß von äußeren oder inneren Zwänge. Falls letztere vorhanden sind, kann dieses Erklärungsmodell natürlich nicht angewandt werden, und es muß z.B. zu psychologisch-psychoanalytischen Erklärungen (innerer Zwang) oder z.B. zu simplen physikalischen Erklärungen (äußerer Zwang) gegriffen werden. Auf den in der Handlungstheorie vieldiskutierten Streit, weil welchem Ausmaß von 'Zwang' man ein Verhalten überhaupt noch als Handlung bezeichnen kann - was auch für die Jurisdiktion relevant ist - können wir hier nicht eingehen.

Auch in der Mathematik, und ebenso in den präskriptiven Disziplinen - die aufgrund unserer Ausführungen in Kap. 1 einen Sonderstatus besitzen - kommen DD-Schemata vor, die hier allerdings nur als deduktive Begründungen fungieren können - denn weder mathematische Sätze noch präskriptive Sätze drücken Ereignisse bzw. konkrete Tatsachen aus, und nur letztere sind voraussagbar oder (kausal) erklärbar im engen Sinne. Ein typisches mathematisches Begründungsmuster ist der induktive Beweis:

Gegeben eine Abzählung (Indizierung) der Objekte $(x_i)_{i \in \mathbb{N}}$ des Objektbereichs, und sei $A(x_i)$ eine beliebig komplexe Eigenschaft von x_i . Das Induktionsschema lautet dann (in der einfachsten Formulierung):

$A(x_1)$

$\forall i \in \mathbb{N} (A(x_i) \rightarrow A(x_{i+1}))$

(Daher:) $\forall i \in \mathbb{N} (A(x_i))$

Die meisten mathematischen Begründungsschemata haben folgende Form:

G: Allgemeinere mathematische Theorie (bzw. Struktur)

A :Einschränkende Bedingung

E: Spezifischere mathematische Theorie (Struktur)

- z.B. läßt sich aus der allgemeinen Theorie der gekrümmten Räume der euklidische Raum als Spezialfall gewinnen (usw.).

In den normativen Disziplinen werden oft speziellere präskriptive Sätze aufgrund allgemeinerer präskriptiver Sätze und präskriptiver Anfangsbedingungen oder aufgrund bedingter Normen (gemischter Sätze) und deskriptiver Anfangsbedingungen begründet - S steht hier für einen Sollens- oder Wertoperator:

G: $A \rightarrow S(B)$

A: A

E: S(B)

Ein Beispiel nach Weingartner (1978, S.118):

G: Für alle Personen x, t und Zeiten t : Wenn sich x zu t in einer Situation befindet, in der sein Leben in Gefahr ist, und wenn es y zu t möglich ist, durch Handlung H das Leben von x zu schützen, dann soll y zu t H durchführen.

Peter liegt zu t im brennenden Wagen und sein Leben ist in Gefahr.

Paul ist es zu t möglich, durch das Herausziehen von Peter aus dem Wagen Peters Leben zu schützen.

(Daher:) Paul soll zu t Peter aus dem brennenden Wagen herausziehen.

3.5 Literaturhinweise:

Zu Erklärung, aber auch anderen Fragen: Hempel, C. G. (1977): Aspekte wissenschaftlicher Erklärung, de Gruyter, Berlin. Ferner auch der Band I der Stegmüller-Reihe (s. Kap. 1) sowie Schurz (1983). Zur Bestätigungsproblematik (vorrangig für deterministische Gesetze). W. Lenzen (1974): Theorien der Bestätigung wissenschaftlicher Hypothesen, frommann, holtzboog. Für den statistischen Fall siehe wieder Bortz (Kap. 2). Vgl. auch: Huff, D. (1956), Wie lügt man mit Statistik, Sanssouci, Zürich. Zur Korrelationsproblematik empfehlenswert: Hummell, H.J., Ziegler, E. (1976): Korrelation und Kausalität, F. Enke, Stuttgart. Über den aktuellen Stand der Erklärungsdebatte informiert: Schurz, G. (Hg., 1988): Erklären und Verstehen in der Wissenschaft (Mit Beiträgen von B.v.Fraassen, P. Gärdenfors, R. Tuomela, M. Friedman, P. Kitcher, G.Schurz. K. Lambert), Oldenbourg (scientia nova), München-Wien. Ein Klassiker des Gesetzesartigkeitsproblems ist: N. Goodman (1975): Tatsache, Fiktion und Voraussage, Suhrkamp, Frankfurt/M. Ein Klassiker zu Erklärungen in den Geschichtswissenschaften ist W. Dray (1957), Laws and Explanations in History, Oxford (Univ. Press).

4. Wissenschaftliche Theorien

4.1 Beobachtungsbegriffe, theoretische Begriffe und Zuordnungsgesetze.

Singuläre Sätze, in denen nur Beobachtungsbegriffe vorkommen, nannten wir allgemein Beobachtungssätze. Ein Satz kann nun aus zwei Gründen kein Beobachtungssatz sein: erstens weil er einen Allquantor enthält und somit eine hypothetische Verallgemeinerung von Beobachtungsdaten darstellt, und zweitens weil er enthält nichtbeobachtbare Begriffe enthält. Mit der Beziehung zwischen singulären empirischen Daten und deren Verallgemeinerungen beschäftigten wir uns im Kap. 3. Durch induktive Verallgemeinerung empirischer Daten kann man jedoch immer nur zu empirischen Gesetzen, nie aber zu theoretischen Gesetzen gelangen. Denn letztere enthalten theoretische Begriffe, welche in ihrem Bedeutungsgehalt das Beobachtbare übersteigen (d.h. nicht durch Beobachtungsbegriffe definierbar sind, s.u.). Der Wissenschaftler gelangt zu solchen theoretischen Begriffen aufgrund kreativer Phantasie, nicht aber durch einfache empirische Induktion. In diesem Kapitel beschäftigen wir uns mit der Beziehung zwischen Beobachtungssätzen und Theorien.

Zunächst: was ist eigentlich ein Beobachtungsbegriff? Man kann zunächst folgende einfache Charakterisierung geben:

B ist ein Beobachtungsbegriff wenn das Zutreffen von B auf einen ebenfalls beobachtbaren Gegenstand von jeder Person unter Normalbedingungen der Beobachtung durch einen (oder mehreren) einfachen Wahrnehmungsakt(en) festgestellt werden kann (wobei wir uns primär auf visuelle Wahrnehmungen, nur sekundär auf die restlichen äußeren Sinne, stützen).

Ob ein Gegenstand beobachtbar ist, kann durch unser sinnesphysiologisches Wissen entschieden werden; ob eine Behauptung über einen beobachtbaren Gegenstand tatsächlich durch einen einfachen Wahrnehmungsakt zustandekam, kann einesteils durch experimentelle Tests - z.B. nonverbale Lernexperimente mit theoretisch unterschiedlich vorbelasteten Personen - entschieden werden, andernteils ist es wiederum Sache bewährten kognitionspsychologischen Wissens.

Es wurde immer wieder bezweifelt, ob es wirklich vollkommen "theoriefreie" Beobachtungssätze gibt. Natürlich enthalten unsere Wahrnehmungsorgane sozusagen "biologisch angeborene" Theorien, z.B. die Verrechnung vom 2dimensionalen Bild auf der Netzhaut zum 3dimensionalen mentalen Sehbild. -- Aber wir verstehen unter Theorie nur bewußte verbale wissenschaftliche Theorien.

Daß Beobachtungssätze in hohem Maß intersubjektiv und daher theorieu-

nabhängig sind, dafür haben wir schon in Kap. 1 argumentiert. Oft wird eingewandt, daß in Beobachtungssätze zumindest der indirekte Einfluß der Sprache miteingeht. Man kann jedoch die Beobachtbarkeit auch in nonverbalen Experimenten prüfen - durch ostensive Lernexperimente, wo nur ja-nein Antworten gegeben werden. Die Idee, Beobachtbarkeit auf diese Weise zu charakterisieren, stammt von Quine. Auf diese Weise kann man zeigen, daß z.B. die Zuni-Indianer, obwohl sie in ihrer Muttersprache keinen Farbbegriff für "Orange" besitzen - durchaus in der Lage sind, den Farbbegriffe "Orange" zu lernen, d.h. sie sind in der Lage, das Merkmal Orange zu beobachten.

Die wichtigsten Merkmale von Beobachtungsbegriffen sind zusammengefaßt folgende:

- *Theorie-Unabhängigkeit*
- *Kultur-Unabhängigkeit*
- *Intersubjektivität*
- *Sprach-Unabhängigkeit*
- *(leichte) ostensive Erlernbarkeit*

Beobachtungssätze in diesem Sinn sind z.B.: "Dies ist jetzt ein Tisch", "Dies ist jetzt rot", "Dieser Zeiger dort bewegt sich jetzt", bis hin zu komplexen visuellen Daten "Dies ist eine Landschaft mit dem und dem Aussehen (evtl. unterstützt durch eine Bilddarstellung)" (dabei sei vorausgesetzt, daß die Ausdrücke "jetzt" und "hier" abkürzend für die entsprechenden Zeit- und Ortsangaben stehen). Die Wahrheit dieser Sätze kann durch visuelle Wahrnehmung festgestellt werden. Natürlich sind auch Beobachtungssätze nicht 100% irrtumssicher - es gibt diverse Fehlerquellen. Aber das behaupten wir nicht (s. der Fallibilismus von Kap. 1.3). Wir behaupten lediglich, daß Beobachtungssätze in obigem Sinn unabhängig von der theoretischen Vorbelastetheit der Person festgestellt werden können, und unter anderem deshalb intersubjektiv sind, und eben daher schließlich wesentlich sicherer sind als (nicht logisch wahre) Gesetze oder Theorien, weshalb (im Regelfall) die letzteren aufgrund der ersteren zu überprüfen sind, und nicht etwa umgekehrt.

Nur wenige Begriffe sind Beobachtungsbegriffe im obigen Sinn. Die niederste Stufe von nicht mehr beobachtbaren Begriffen sind die *Dispositionsbegriffe* wie z.B. "x ist wasserlöslich". Der Satz "diese Substanz ist wasserlöslich" kann nicht mehr durch einen einfachen Wahrnehmungsakt überprüft werden. Es muß dazu ein *Experiment bzw. ein Test* durchgeführt werden, in dem der zu beobachtende Gegenstand x einer empirischen Transformation unterworfen wird: er wird ins Wasser gegeben. Aufgrund des beobachtbaren Experimentresultats ("x löst sich auf/nicht auf") kann dann erst über die Löslichkeit bzw. Nicht-Löslichkeit von x entschieden werden. Wie Carnap gezeigt hat, lassen sich derartige

Dispositionsbegriffe nicht mehr vollständig über Definitionen auf Beobachtungsbegriffe zurückführen. Es sind nur mehr unvollständige Zurückführungen auf Beobachtungsbegriffe in Form von *Zuordnungsgesetzen* möglich (auch Reduktionssätze genannt), die im einfachsten Fall folgende Form haben:

$$Tx \rightarrow (Dx \leftrightarrow Bx)$$

Wenn x Testbedingung T erfüllt, dann: (x hat Dispositionsmerkmal D g.d.w. x das beobachtbare Verhalten B zeigt).

z.B.:

Wenn x in Wasser gegeben, dann: (x ist wasserlöslich \leftrightarrow x löst sich auf)

T und B sind hier Beobachtungsbegriffe, und D ist der nichtbeobachtbare Dispositionsbegriff. Nur unter der speziellen Testbedingung Tx ist Dx mit Bx äquivalent. Falls jedoch Tx nicht erfüllt ist, d.h. in unserem Beispiel falls x *nicht* ins Wasser gegeben wurde, kann über Dx durch Beobachtung nichts ausgesagt werden. Das ist auch der Grund, warum Dx nicht durch Beobachtungsbegriffe definierbar ist.

Im Unterschied zu einfachen Dispositionsbegriffen wie "wasserlöslich", deren Bedeutung bereits durch ein einzelnes Zuordnungsgesetz hinreichend expliziert und auch erschöpft ist, ist es das wesentliche Merkmal von theoretischen Begriffen wissenschaftlicher Theorien, *durch eine Mehrzahl von Zuordnungsgesetzen mithilfe ganz unterschiedlicher Testbedingungen empirisch charakterisiert zu werden*. Dabei ist ein Begriff in der Regel umso theoretisch höherliegender, je mehr unterschiedliche Zuordnungsgesetze ihn charakterisieren. Theoretisch schon höherstehende Begriffe als die vorhingenannten Dispositionsbegriffe sind Substanztypbezeichnungen, z.B. "Sauerstoffgas". Betrachten wir die zwei Zuordnungsgesetze:

(Z1) (Oxidbildung) Wenn x in ein abgeschlossenes Gefäß mit Kohlenstoff bei hoher Temperatur eingeleitet wird, dann: x ist Sauerstoffgas g.d.w. der Kohlenstoff unter heller Flamme verbrennt; formal: $T1x \rightarrow (Sx \leftrightarrow B1x)$.

(Z2) (Knallgasreaktion) Wenn x mit Wasserstoff vereinigt und entzunden wird, dann: x ist Sauerstoffgas g.d.w. das Gemisch explodiert und daraus Wasser entsteht; formal: $T2 \rightarrow (Sx \leftrightarrow B2x)$

Es gibt noch viele weitere Zuordnungsgesetze für Sauerstoffgas. An diesem Beispiel erkennt man deutlich, warum Zuordnungsgesetze *nicht als analytisch bzw. definitorisch wahr*, also nicht als bloße Begriffskonventionen ohne empirischen Gehalt aufgefaßt werden dürfen. Die Zuordnungsgesetze Z1 und Z2 zusammen haben

folgendes empirisches Gesetz zur logischen Konsequenz:

(EK) $\forall x(T2x \wedge B2x \wedge T1x \rightarrow B1x)$

Jedes Gas x, das mit Wasserstoff zusammengefügt und entzunden unter Explosion Wasser ergibt, wird - bei hoher Temperatur in ein geschlossenes Gefäß mit Kohlenstoff geleitet - dessen Verbrennung mit heller Flamme bewirken.

Man darf also niemals mehrere Zuordnungsgesetze zusammen als analytisch bezeichnen, denn schon zwei zusammen haben empirischen Gehalt. Man dürfte maximal eine unter allen Zuordnungsgesetzen als analytisch im Gegensatz zu den anderen auszeichnen - aber das wäre reine Willkür, weil es lediglich dieses eine Zuordnungsgesetz 'immun' gegenüber möglichen Korrekturen oder Revisionen machen würde. Wir sehen also: *Zuordnungsgesetze sind, wie alle anderen Bestandteile von Theorien, hypothetischer Natur, also bewährbar und korrigierbar (falsifizierbar).*

Daß theoretische Begriffe nicht definierbar sind, und Zuordnungsgesetze nicht analytisch, sondern indirekt empirisch überprüfbar, war eine der einschneidendsten wissenschaftstheoretischen Entdeckungen. Der reduktionistische (bzw. naive) Empirismus, der noch im Wiener Kreis stark vertreten war - demzufolge alle wissenschaftlichen Begriffe durch Beobachtungsbegriffe definierbar sein sollten - wurde dadurch klar widerlegt. Die heutige "postpositivistische" Wissenschaftstheorie ist vom früheren naiven Empirismus abgegangen; vertretbar ist nur mehr unser "minimaler" Empirismus von Kap. 1.

Zuordnungsgesetze bei quantitativen Begriffen sind häufig das, was man *Meßgesetze* nennt. Erläutern wir dies am physikalisch-theoretischen Begriff der Masse. Man kann Masse auf unterschiedliche Arten messen, z.B. auf der Balkenwaage, oder auf der Federwaage:

(Z1) $T_1x \rightarrow (Mx \leftrightarrow R_1x)$
 x auf Balkenwaage $m(x) = \text{so und so Zeigerausschlag so und so}$

(Z2) $T_2x \rightarrow (Mx \leftrightarrow R_2x)$
 x auf Federwaage $m(x) = \text{so und so Zeigerausschlag so und so}$

Balkenwaage-Messung und Federwaage-Messung beruhen auf unterschiedlichen physikalischen Mechanismen. Z.B. ändert sich das Ergebnis der Balkenwaage-Messung auch dann nicht, wenn die Schwerkraft sich ändert - z.B. am Mond. Das Ergebnis der Federwaage ändert sich schon. Wieder haben beide Zuordnungsgesetze zusammen einen empirischen Gehalt, nämlich: wenn ein Gegenstand x auf einer

Balkenwaage 1 kg genau aufwiegt, dann wird er auf einer Federwaage die Feder so und so weit zusammendrücken.

Beide Messungen messen die "schwere" Masse, setzen also eine wirkende Gravitationskraft voraus; beide wären im Vakuum z.B. nicht anwendbar. Hier kann man die Masse z.B. durch Stoßexperimente mittels dem Impulssatz bestimmt, man nennt das die sogenannte "träge" Masse, die mit der "schweren" identisch ist - wieder ergeben sich neue empirische Konsequenzen.

Auch in den Human-und Sozialwissenschaften sind die meisten Begriffe theoretischer Natur. Die empirische und theoretische Begriffe verbindenden Zuordnungsgesetze heißen hier *Indikatorbeziehungen* bzw. Indikatorgesetze. Auch sie sind prinzipiell hypothetischer Natur, und nicht definitiv wahr (was in manchen Lehrbüchern heute leider immer noch behauptet wird).

Einige Beispiele:

Theoretischer Begriffe:

Einstellung von Personen

Indikatoren:

Antwort auf eine Frage eines Interviews zu bestimmten Fragen

Kognitives Niveau eines Kindes

Resultat in einem kognitiven Leistungstest.

Charaktertendenz von Personen, z.B. Aggressivität, emotionelle Labilität

Verhalten in gewissen experimentell vorgegebenen Situationen; oder Resultat eines Persönlichkeitstests, etc.

Die politische Einstellung einer sozialen Schicht

Z.B. ihr durchschnittliches Wählerverhalten; oder durchschnittliches

Umfrageergebnis

Die logische Form dieser Indikatorbeziehungen ist dieselbe: z.B. "Wenn Person x einem Einstellungsinterview I ausgesetzt wird, dann: x hat Einstellung E g.d.w. x Antworten A gibt" (usw.). Der Unterschied zu den Naturwissenschaften besteht darin, daß diese Indikatorbeziehungen wesentlich willkürlicher und weniger durch die Theorie vorgegeben sind. Dennoch ist dieser Unterschied nur ein gradueller, kein prinzipieller.

Es sei bemerkt, daß Zuordnungsgesetze nicht immer obige einfache Form haben; sie sind oft auch nicht in den Axiomen der Theorie direkt enthalten, sondern ergeben sich als logische Konsequenzen. Im Fall von quantitativen Theorien sind Zuordnungsgesetze zugleich Meßgesetze. (s.o.). *Als Zuordnungsgesetz im weiten Sinn kann man jedes Axiom oder Theorem einer Theorie auffassen, welches sowohl theoretische wie empirische Begriffe enthält.* - Neben Zuordnungsgesetzen und empirischen Konsequenzen sind ein weiterer wichtiger Theorienbestandteil rein theoretische Gesetze, die nur theoretische Begriffe enthalten. Ein rein theoretisches Gesetz unserer obigen "Sauerstofftheorie" wäre z.B.

Für alle x: x ist Sauerstoff g.d.w. x aus homonuklearen Molekülen des Atoms 0816 besteht (i.e. 8 Protonen und 8 Neutronen im Kern, 8 Elektronen in der Hülle)

Ein rein theoretisches Gesetz (Korrelationsgesetz) aus der Piagetschen Entwicklungspsychologie wäre

Kinder mit hohem kognitiven Niveau weisen auch ein hohes Niveau sozialer Reife auf.

- sowohl "kognitives Niveau" wie "soziale Reife" sind theoretische Merkmale.

4.2 Exkurs: Erkenntnistheoretische Hierarchie von Begriffen

Genau genommen sind auch die Begriffe, die wir üblicherweise Beobachtungsbegriffe nennen, eigentlich Dispositionsbegriffe, sozusagen "Dispositionen 0.ter Stufe", insofern wir mit ihnen realistische Hypothesen verbinden. Ein Beispiel. Wenn ich beobachte, dieser Baum dort ist braun, so drücke ich damit nicht nur mein momentanes Beobachtungserlebnis aus, sondern darüberhinaus, daß es den Baum unabhängig von meinem Beobachtungserlebnis wirklich gibt, und daß er unabhängig davon, ob ich ihn beobachte, braun ist. Also, daß er z.B. auch dann braun ist, wenn ich wieder wegsehe. Ich schließe also ebenfalls von einem "Test"- meinem momentanen Wahrnehmungseindruck - auf ein Dispositionsmerkmal des Baumes, seine Farbe, die er unabhängig von meinen Wahrnehmungen hat.

Das hat offenbar dieselbe Form wie ein Zuordnungsgesetz in obigem Sinne, nur daß die Testbedingungen hierbei nicht darin bestehen, den Gegenstand einer bestimmten experimentellen Operation zu unterwerfen (wie das Auflösen des Zuckers in Wasser), sondern die Testbedingungen fallen hier einfach mit den Normalbedingungen der Beobachtung zusammen.

Begriffe, die sich rein auf Beobachtungserlebnisse beziehen, wollen wir *phänomenale Begriffe* nennen. Beobachtungsbegriffe wie "ist braun" im üblichen alltäglichen Sinn sind dagegen *Realbegriffe*. Der Schluß von phänomenalen Begriffen auf Realbegriffe, bzw. von phänomenalen Beobachtungssätzen auf Realbeobachtungssätze wird vorallem in der Erkenntnistheorie behandelt. Phänomenale Sätze nennt man auch *introspektiv* - sie sind insofern "unkorrigierbar", weil der introspektive Satz "ich habe jetzt eine Braunempfindung" auch dann wahr bleibt, wenn dort tatsächlich kein brauner Baum steht, sondern ich einen solchen nur *halluziniere*.

Es ergibt sich folgende *Hierarchie von Begriffen*:

Theoretische Begriffe (Dispositionsbegriffe 2. Stufe)	Viele Indikatoren (Zuordnungsgesetze) Ladung, Kraft, Persönlichkeit
Dispositionsbegriffe (1. Stufe)	Nur ein Indikator (Zuordnungsgesetz) Wasserlöslich
Beobachtungsbegriffe realistischen Sinn) (Dispositionsbegriffe 0.ter Stufe)	Testbedingung = Normalbedingungen (im der Wahrnehmung rot, länger als, rund
Phänomenale Begriffe	"mir erscheint so und so" Introspektive Merkmale

4.3. Strukturelle Bestandteile und methodologische Merkmale wissenschaftlicher Theorien - am Beispiel der Newtonischen Physik

Wir fassen die wichtigsten strukturellen Bestandteile wissenschaftlicher Theorien noch einmal zusammen.

1. In der Sprache unterscheiden wir zwischen empirischen und theoretischen Begriffen - allgemeiner, zwischen der *rein empirischen Subsprache* und der *theoretischen Gesamtsprache*.

2. In der Theorie unterscheiden wir zum einem zwischen *Axiomen* der Theorie, und Theoremen oder *Konsequenzen* - das sind jene Sätze, die aus der Theorie logisch

folgen.

Die Menge aller Konsequenzen nennen wir den *Gesamtgehalt* der Theorie.

Sowohl unter den Axiomen wie unter den Konsequenzen unterscheiden wir wie folgt:

2.1 *Rein theoretische Gesetze* - das sind solche Sätze des Gesamtgehalts, die neben logischen Begriffen nur theoretische Begriffe enthalten.

2.2 *Zuordnungsgesetze* - das sind solche Sätze des Gesamtgehalts, die neben logischen Begriffen sowohl empirische wie theoretische Begriffe enthalten.

2.3 *Empirische Konsequenzen* - jene Sätze des Gesamtgehalts, die neben logischen nur empirische Begriffe enthalten. Die Menge der empirischen Konsequenzen nennt man den *empirischen Gehalt* einer Theorie.

Zumeist bestehen die Axiome einer Theorie aus theoretischen Gesetzen und Zuordnungsgesetzen. Zuordnungsgesetze finden sich meistens sowohl in den Axiomen wie in den Theoremen (Konsequenzen). Die empirischen Konsequenzen finden sich im Regelfall nur unter den Konsequenzen der Theorie, nicht in den Axiomen.

Eine weitere wichtige Unterscheidung ist die zwischen *Kern* und *Peripherie* von Theorien. Der Theoriekern umfaßt jene Gesetze - theoretische Gesetze oder Zuordnungsgesetze - die für die Theorie von entscheidender Bedeutung sind, die sozusagen ihre Identität definieren. Ändert man den Kern, so ist man zu einer anderen Theorie übergegangen.

Die Peripherie umfaßt dagegen speziellere Gesetzhypothesen der Theorie, z.B. spezielle Randbedingungen oder Zusatzannahmen. Ändert man die Peripherie einer Theorie, so hat man die aktuelle *Theorieversion* geändert, man ist jedoch immer noch innerhalb des alten theoretischen "Rahmens" geblieben.

Zwischen Kern und Peripherie besteht eine graduelle Beziehung, eine Abstufung.

Wir erläutern beides an einer quantitativen Theorie, der Newtonschen Mechanik.

Hier sind Ort und Zeit, bzw. Geschwindigkeit und Beschleunigung die Beobachtungsbegriffe. Masse und Kraft sind dagegen theoretische Begriffe. Als Kern der Newtonschen Mechanik wird zumeist das Axiom

(N1) Kraft = Masse mal Beschleunigung

angeführt (welches sowohl theoretische wie empirische Begriffe enthält). Dieses Axiom allein hat jedoch (wie man beweisen kann) noch *überhaupt keinen empirischen Gehalt* - Kraft und Masse sind nicht beobachtbar. Nehmen wir das andere Newtonsche Axiom

(N2) Kraft = Gegenkraft, oder: Summe aller Kräfte = Null.

hinzu. Es enthält nur theoretische Begriffe, ist also ein rein theoretisches Gesetz. Aus N1 und N2 läßt sich - zusammen mit der Definition des Drehmomentes - das Hebelgesetz der Balkenwaage herleiten, woraus ein Zuordnungsgesetz (i.e. ein empirisches Meßgesetz) für die Masse folgt. Aber der Kraftbegriff ist immer noch unbestimmt. Auch N1 und N2 zusammen haben sozusagen noch "fast keinen" empirischen Gehalt.

Der entscheidende empirische Gehalt kommt nun erst durch Zuordnungsgesetze für verschiedene Kraftarten zustande, z.B.:

(N3) Gravitationskraft = Funktion von: Massen der beiden Körper und ihr Abstand zum Quadrat.

(N1)-(N3) ergeben zusammen eine Differentialgleichung, deren (exakte oder approximative) Lösung es ermöglicht, die Bewegung von Körpern in Schwerfeldern vorauszusagen, in Abhängigkeiten von den anwesenden Körpern, die eine nichtvernachlässigbare Gravitationskraft ausüben, sowie von ihren Anfangs- und Randbedingungen. Dadurch ergibt sich eine Vielzahl von ganz unterschiedlichen, erklärbaren bzw. voraussagbaren empirischen Phänomenen, z.B.

a) elliptische Planetenbewegungen - im Schwerfeld der Sonne, dadurch indirekt Jahreszeiten usw.;

b) Bahnen von Körpern im Weltraum im Schwerfeld der Erde; Bewegung des Mondes, indirekt z.B. auch von Ebbe und Flut, von Meteoren oder Kometen, Raketen; zugleich aber auch

c) die Bewegungen von Körpern an der Erdoberfläche bei relativ konstanter Schwerkraft., z.B. Wurfgeschosse, freier Fall, Bewegungen auf schiefer Ebene, usw.

Die Tatsache, daß plötzlich eine Vielzahl unterschiedlicher Phänomene erklärbar werden (=Systemcharakter und empirische Kreativität), die ihrerseits ganz unterschiedlichen Typs sind (=Globalität) und daß insbesondere die - im Aristotelischen Weltbild zuvor getrennte - Physik der Erdkörper und die der Himmelskörper nun unter eine gemeinsame Theorie subsumierbar werden (=Vereinheitlichungskraft), machte den durchschlagenden historischen Erfolg der Newtonschen Theorie aus.

Damit können wir die *methodologischen Merkmale (guter) wissenschaftlicher Theorien* so zusammenfassen:

a) *der Systemcharakter oder 'Holismus' ihres empirischen Gehalts*: ein isoliertes

theoretisches Gesetz einer Theorie hat noch keinen empirischen Gehalt, mehrere zusammen können aber einen enormen empirischen Gehalt haben

b) damit zusammenhängend die *empirische Kreativität von Theorien* - ihr empirischer Gehalt ist offen und kann durch Hinzufügung neuer Spezialgesetze unvorhersehbar vermehrt werden

c) die *Globalität* von Theorien - die empirische Erklärungskraft erfaßt empirische Phänomene ganz unterschiedlichen Typs, wodurch

d) Theorien eine hohe *Vereinheitlichungskraft* für das empirische Wissen besitzen.

4.4 Holismus der Theorienüberprüfung - am Beispiel der Newtonischen Physik

Die empirische Überprüfung von wissenschaftlichen Theorien vollzieht sich über die Überprüfung ihres empirischen Gehaltes. Da ein einzelnes theoretisches Gesetz noch keinen empirischen Gehalt hat, kann man ein einzelnes theoretisches Gesetz auch nicht empirisch prüfen - vielmehr läßt sich nur ein ganzes *System von theoretischen Gesetzen plus Zuordnungsgesetzen plus Hilshypothesen überprüfen*, das stark genug ist, um interessante empirische Voraussagen oder Erklärungen zu liefern. Betrachten wir das Beispiel der Voraussage (Erklärung) von Planetenbahnen mittels der Newtonschen Theorie:

(K1) Kraft=Masse mal Beschleunigung

(K2) Aktio=Reaktio

(S3) Gravitationskraft = Funktion von: Massen der beiden Körper und Abstand zum Quadrat.

(H1) (Hilfshypothese) Auf Planet a wirken im relevanten Zeitintervall die und die Gravitationskräfte

P: Bewegungsbahn von a = Funktion von: Anfangsbedingungen, Zeit.

P ist das Explanandum bzw. Prognosezendum - also ein empirischer Satz, der aus den Gesetze K1, K2, S1 plus Hilfshypothese (H1) logisch (bzw. mathematisch) folgt. K1 und K2 sind *Kerngesetze*, gehören zum Kern der Newtonischen Theorie. S1 ist ein spezielles Kraftgesetz, sozusagen in der Mitte von Kern und Peripherie. H1 ist eine Hilfshypothese, sie liegt ganz außen in der Peripherie, ist als für den Kern eher unwichtig.

Stimmt das Explanandum tatsächlich mit der Beobachtung überein, so ist der ganze Komplex K1+K2+S1+H1 *zusammen bewährt*. Angenommen aber, die empirische Prognose E stimmt nicht mit den tatsächlichen Beobachtungen überein - was dann? Dann entsteht die typische Situation, die man (seit Pierre Duhem) den *Holismus der Theorienfalsifikation* nennt: der Gesamtkomplex K1+K2+S1+H1 ist dann falsifiziert, und man weiß, daß irgendeiner dieser Prämissen (oder nur ein Prämissenteil) falsch sein muß, aber man weiß nicht, *welcher Teil falsch ist*.

In der holistischen Falsifikationssituation ist es wichtig, daß die Menge der Prämissen der Theorie graduell eingeteilt ist von Kernbereich zu Peripheriebereich. Der Kernbereich - er entspricht ungefähr dem, was Kuhn unter Paradigma versteht - enthält die fundamentalen Axiome, d.h. würde man diese Bestandteile preisgeben, so würde man nicht mehr von einer neuen Variante derselben Theorie sprechen, sondern von einer neuen Theorie. Der periphere Bereich enthält Spezialgesetze und vorallem Hilfshypothesen; durch Änderung des peripheren Bereichs erhält man neue Varianten derselben Theorie. Z.B. ist "Kraft = Masse mal Beschleunigung" ein Kernbestandteil der Newtonschen Theorie, dagegen obige Hilfshypothese (H1) eindeutig ein peripherer Bestandteil. Oder "es gibt unbewußte Wünsche" ist ein Kernbestandteil der Freudschen Theorie, dagegen seine Todestriebshypothese eher ein peripherer Bestandteil (man kann die Psychoanalyse auch ohne Todestriebshypothese entwickeln).

Wenn nun die empirische Prognose eines Theorienkomplexes in Konflikt mit der Beobachtung gerät, wird man zuerst den Theoriekern beibehalten und periphere Bestandteile ändern. In obigem Beispiel würde man als erstes die Hilfshypothese anzweifeln und vermuten, daß man einen relevanten Gravitationseinfluß auf den Planeten a durch einen bisher unbekanntem Nachbarplaneten oder sonstigen

Himmelskörper übersehen hat. Eine solche Hypothese, die eine Änderung in der Theorieperipherie nur zu dem Zweck bewirkt, damit die Theorie mit einem widerspenstigen Datum wieder vereinbar wird, nennt man auch *ad hoc Hypothese*. Eine ad hoc-Hypothese ist jedoch an sich noch nichts Schlechtes - alles hängt davon ab, ob sie später unabhängige empirische Bewährung erfährt oder nicht.

Folgende zwei wissenschaftsgeschichtliche Beispiele demonstrieren dies: (A) Als man etwa um 1920 - damals waren nur 6 Planeten bekannt - entdeckte, daß die Planetenbahn des Uran von der durch die Newtonische Theorie vorausgesagten Bahn abwich, postulierten Adams und Leverrier ad hoc die Existenz eines neuen Planeten, des Neptun, dessen Gravitationswirkung die Abweichungen erklären sollte. Diese Planet konnte später unabhängig empirisch bestätigt werden (durch stärkere Teleskope), die veränderte Hilfshypothese und somit die gesamte neue Version der Newtonschen Himmelsmechanik hat sich unabhängig empirisch bewährt.

(B) Ebenso postulierte Leverrier für die Abweichung der Planetenbahn des Merkurs von den Newtonischen Prognosen die Existenz eines neuen Planeten, genannt Vulkan, der jedoch in der Folgezeit hartnäckig teleskopisch nicht auffindbar war und auch sonst nicht unabhängig bestätigt werden konnte. Die ad hoc-Hypothese des Vulkan und damit indirekt der Newtonische Theoriekern wurde dadurch geschwächt; in der Tat zeigte sich später, daß die Erklärung der Abweichung erst mit der Einsteinschen Relativitätstheorie - also einem neuen Theoriekern (!) - möglich war.

Wie Kuhn (1976) gezeigt hat, ist es zunächst immer möglich, den Kern einer Theorie gegenüber zunächst widerspenstiger Erfahrung durch solche ad hoc-Abänderungen zu retten. Solche ad hoc-Änderungen sind jedoch, wie Popper und Lakatos (1974) als methodologische Regel vorschlugen, nur dann wissenschaftlich legitim, solange sie zu einer neuen Theorieversion mit größerem oder *zumindest mit nicht geringerem empirischen Gehalt* führen. Wenn sich die widersprechenden Erfahrungen mehren und die zu ihrer Beseitigung nötigen ad hoc-Hypothesen immer komplizierter und künstlicher werden, ohne neuen empirischen Gehalt zu erzeugen, erst recht wenn dadurch der empirische Gehalt sogar nach und nach immer mehr eingeschränkt wird, dann wird man am alten Theoriekern zweifeln und es mit einem neuen Theoriekern versuchen. Erst dann, wenn man einen neuen Theoriekern gefunden hat, der erfolgreicher ist als der bisherige (wie z.B. Einsteins Relativitätstheorie), wird man den bisherigen Theoriekern als 'falsifiziert' ansehen.

4. 5 Zweites Beispiel: Piagets kognitive Entwicklungspsychologie

Die zentrale These von Piagets kognitiver Entwicklungspsychologie besagt, daß die kindliche Intelligenzentwicklung auf der *stufenförmigen Ausbildung von allgemeinen logisch - strukturellen Fähigkeiten* beruht. Sind diese einmal ausgebildet, so können

sie innerhalb kurzer Zeit sozusagen überall angewendet werden. Vorher dagegen werden sie einfach "nicht begriffen". Eine weitere Kernthese besagt, daß die sogenannte *konkret-operationelle Stufe* vom Kind mit ca. 6-7 Jahren erreicht wird. - Daneben gibt es noch die *sensomotorische Stufe* (mit 2 Jahren), und die *formal-operationale Stufe* (mit 13-14 Jahren). - Die konkret-operationelle Stufe ist durch die Ausbildung folgender Fähigkeiten gekennzeichnet: Erstens der *Perspektivenwechsel*, d.h. die Fähigkeit, die eigene Sehperspektive von jener anderer Personen zu unterscheiden, und damit zusammenhängend der Übergang vom egozentrischen zum soziozentrischen Denken. Zweitens die Fähigkeit zu erkennen, daß bestimmte Operationen, die das Aussehen von bestimmten Dingen verändern, auch wieder rückgängig gemacht werden können - die sogenannte *Reversibilität* dieser Operationen - und damit zusammenhängend die Einsicht, daß bestimmte Eigenschaften der Dinge unter solchen Operationen invariant erhalten bleiben - die Fähigkeit zur *Invariantenerkennung*. Reversibilität und Invariantenerkennung untersuchte Piaget vorallem in zwei Bereichen. Einerseits die *Invarianz der Anzahl* bzw. Zahl von Gegenständen unter ihren verschiedenen Anordnungsmöglichkeiten. Andererseits die *Invarianz der Menge* einer gewissen Substanz unter ihren verschiedenen Verformungsmöglichkeiten.

Alle diese Fähigkeiten hat Piaget weiters an gewissen *Tests* überprüft. Die Fähigkeit des Perspektivenwechsels anhand eines dreidimensionalen Modells einer *Gebirgslandschaft*; die Kinder saßen an vier verschiedenen Seiten und mußten anhand von vorgezeigten Bildern sagen, wie das Gebirge von der bzw. der Position aussieht. Die Invarianz der Zahl wurde mit einem Zahltest getestet, wo eine bestimmte Zahl von Münzen auf Bildern einmal weit auseinanderliegend angeordnet wurde, das andere Mal eng beisammen, usw. Es wurde jedesmal gefragt "sind auf diesem Bild mehr, gleich viel oder weniger Münzen als auf jenem?" -- Die Invarianz der Substanzmenge wurde schließlich mit Piagets berühmten *Tonkugelexperiment* getestet, wo eine Tonkugel einmal als Kugel präsentiert wurde, und dann zu einer langen Wurst auseinandergezogen wurde, und es wurde gefragt, "ist das nun mehr, oder gleichviel geblieben?". Bekannt sind hier auch die Flüssigkeitsumschüttexperimente, wo eine bestimmte Menge Wasser von einem breiten in ein sehr enges und hohes Glas umgefüllt wird.

bei allen diesen Test ergab sich, daß die entsprechende Fähigkeit vor 6 Jahren überwiegend nicht beherrscht wurde, nach 7 Jahren durchwegs schon; wodurch Piaget seine Stufentheorie der kognitiven Intelligenzentwicklung als bewährt ansah. Eine wichtige Voraussetzung, die Piaget hierfür machen mußte, ist natürlich, daß seine Tests tatsächlich genau die kognitiven Fähigkeiten gemessen hatten, die er damit messen wollte, und nicht zugleich andere kognitive Schwierigkeitsvariablen mitgemessen hatten. Anders gesprochen, daß die Tests wirklich *selektive* (bzw.

adäquate) Tests waren.

Man kann das in folgendem schematischen Argument zusammenfassen:

K1 Die kindliche Intelligenzentwicklung beruht (primär) auf der stufenförmigen Entwicklung von logisch - strukturellen Fähigkeiten.

K2-4 Die konkret-operationelle Stufe ist gekennzeichnet durch die Erlangung der Fähigkeit zu:

- (K2) Perspektivenwechsel
- (K3) Reversibilität von Operationen, und
- (K3) Invariantenerkennung

S1 Mit 6-7 Jahren erreichen die Kinder die konkret-operationelle Stufe

S2-4: (S2) Die Fähigkeit zum Perspektivenwechsel äußert sich empirisch bei den Raumperspektiven des räumlichen Sehens.

Die Fähigkeit zu Reversibilität und Invariantenerkennung äußert sich in der Fähigkeit, (S3) die Invarianz der Anzahl zu erkennen, sowie in der Fähigkeit, (S4) die Invarianz der Substanzmenge zu erkennen.

Z1-3: Selektiver Test für Fähigkeit zum Perspektivenwechsel ist der Piagetsche Gebirgstest.

Selektiver Test für Fähigkeit, die Invarianz der Anzahl zu erkennen, ist der Piagetsche Zahltest.

Selektiver Test für Fähigkeit, die Invarianz der Substanzmenge zu erkennen, ist der Piagetsche Tonkugelttest.

Prognose: Kinder beherrschen (hochsignifikant) die drei Piagetschen Tests vor 6 Jahren nicht, nach 7 Jahren schon.

K1-4 sind Kerngesetze; mit ihnen steht und fällt die Piagetsche Theorie. S1-4 sind Spezialgesetze, immer noch relativ wichtig. Z1-3 sind sehr periphere Zuordnungsgesetze, die die Selektivität bzw. Adäquatheit der Piagetschen Tests betreffen.

Zu Piagets kognitiver Entwicklungspsychologie gab es in den 60ern und 70ern sehr viele kritische Replikations- und Überprüfungsstudien. Man kann an deren Verlauf gut wieder den Holismus der Theorienüberprüfung studieren.

Zunächst wurde speziell beim Sehtest gezweifelt, ob der Piagetsche Sehtest mit

Gebirgslandschaft wirklich selektiv war. Man führte einen viel einfacheren Test durch: mit einer Schachtel, die vier Seiten hatte, deren vier Seiten alle eine andere Farbe hatte. Und fragte dann die Kinder, nachdem man sie ein oder mehrmals um die Schachtel herumgeführt hatte, und dann auf eine bestimmte Seite niedersetzen ließ: "welche Farbe hat die Schachtel, von dieser (anderen) Seite aus betrachtet?" Ergebnis: die Kinder beherrschten schon mit 4 Jahren diesen einfacheren Sehtest zum Perspektivenwechsel.

Daraus schloß man zunächst, daß offenbar Piagets Sehtest nicht selektiv war, weil er nicht nur die prinzipielle Fähigkeit zum Perspektivenwechsel testete, sondern zugleich ein hohes räumliches Vorstellungsvermögen erforderte - schließlich mußten die Kinder ein kompliziertes Gebirge in ihrem Kopf drehen. Man schlug also zunächst vor, die peripheren Gesetze der Piagetschen Theorie zu ändern, nämlich S1 sowie Z1: man vermutete also, daß die konkret-operationelle Stufe doch schon *früher* erreicht wird als von Piaget vermutet, und man änderte das Zuordnungsgesetz für einen adäquaten Perspektivenwechseltest.

Analoges fand man beim Zahltest heraus. Die zusätzliche Schwierigkeitsvariable war hier die Sprache - die Frage, ob die Kinder richtig verstanden, was mit "mehr Münzen" genau gemeint war - mehr im Sinne der Zahl, oder mehr im Sinne der von den Münzen beanspruchten Fläche? Man führte den Zahltest dann mithilfe von nonverbalen Lernexperimenten durch (man belohnte die Kinder, wenn sie Bilder mit der gleichen Anzahl von Münzen darauf einander zuordnen konnten) - und siehe da, der Zahltest wurde dann ebenfalls schon von 4-5 jährigen beherrscht.

Andere Autoren aber brachten Experimente oder Befunde vor, die man nicht mehr mit einer solchen Peripherieänderung erklären konnte. Beispielsweise die *Aussagenlogik*. Gemäß Piaget müßten die aussagenlogischen Gesetze mit 13-14 Jahren, bei Erreichen der formal-operationalen Stufe beherrscht werden. Tatsächlich aber werden gewisse aussagenlogische Gesetze, z.B. Modus Ponens, schon mit drei Jahren beherrscht, andere dagegen, z.B. Modus Tollens, werden von etlichen Kindern bzw. Personen überhaupt nie richtig begriffen. -- Das läßt sich durch eine Änderung von Testgesetzen Ti oder Spezialgesetzen Si nicht mehr erklären, sondern greift die Kernannahmen der Piagetschen Theorie an.

Analoges ergab sich bei der sogenannten *Erhaltung des Objekts* - daß Objekte, die vorher sichtbar waren und nachher unsichtbar geworden sind, nicht plötzlich aufhören zu existieren, sondern weiterexistieren. Eine gewisse Variante dieser Fähigkeit wird nach Piaget bereits nach Vollendung der sensomotorischen Stufe, mit 2 Jahren, beherrscht: wenn man einen Ball hinter einer Wand versteckt, so suchen die Kinder nach dem Ball. Eine andere Variante, nämlich die Auflösung von Zucker in Wasser, wird aber noch nicht einmal mit 6 Jahren beherrscht - man löst sichtbaren Zucker in Wasser auf und fragt dann die Kinder, nachdem die kleinsten

Zuckerkrümeln im Wasser verschwunden sind: "was ist nun mit dem Zucker passiert - ist er nicht mehr da?".

Solche Befunde führten etliche Psychologen dann dazu, die *Kernannahme* K1 der Piagetschen Theorie *preiszugeben*. Die *alternative* Theorie, die entwickelt wurde, besagt, daß die Entwicklung der Intelligenz nicht über die Ausbildung allgemein-abstrakter Strukturen bzw. Fähigkeiten vor sich geht, sondern auf der Entwicklung *inhaltsspezifischer und inhaltsgebunder Fähigkeiten* beruht, welche zunächst nur in einem bestimmten Anwendungsbereich beherrscht werden, und erst später durch Differenzierung und Generalisierung auf andere Bereiche übertragen werden können.

4.6 Adäquatheit von Zuordnungsgesetzen bzw. Indikatoren in den Human- und Sozialwissenschaften

Gerade in diesem Bereich ist es wichtig, die Indikatorbeziehungen zu reflektieren und problematisieren: was bei einer Untersuchung über ein theoretisches Merkmal herauskommt, kann nämlich in hohem Maß davon abhängen, welchen Indikator man für das Merkmal wählt. Es gibt zwei Arten von verzerrten Indikatoren:

Verzerrte Indikatoren durch Einseitigkeit:

Ein Beispiel hierfür waren Piagets kognitive Leistungstests. Es sind jedoch zum theoretischen Merkmal "operationelles Denken" viele *verschiedene* Indikatoren, d.h. verschiedene Testsituationen vom selben strukturellen Typ möglich; und es zeigte sich, daß man bei Wahl anderer Indikatoren (etwa nonverbale Experimente; oder Experimente unter Ausschaltung gewisser Variablen wie Gedächtnis oder Sprachbeherrschung) zu ganz anderen Ergebnissen kam.

Zur Vermeidung von einseitigen Indikatoren ist folgende Regel einzuhalten: *Eine theoretische Hypothese ist nicht nur anhand von einem, sondern anhand so viel als möglichen plausiblen Indikator zu testen*. Gerade wegen der erwähnten relativen 'Willkür' bzw. theoretischen Unterbestimmtheit von Indikatoren in den Human- und Sozialwissenschaften ist die Befolgung dieser Regel wichtig.

Verzerrte Indikatoren durch versteckte Variablen: Wir nehmen hier das Beispiel von Umfrageergebnissen. Häufig werden diese als Indikatoren für die Einstellung von Personen verwendet. Z.B. dient die Frage "wie wichtig sind Ihnen andere Menschen?" als Indikator für Extrovertiertheit, oder "wie wichtig ist für Sie Umweltschutz, im Vergleich zu Wirtschaftswachstum?", als Maß für Ökologiebewußtsein. Was aber in Interviews gemessen wird, ist zunächst nur die verbal geäußerte Einstellung, welche mit der tatsächlichen Einstellung im alltäglichen Verhalten keineswegs übereinstimmen muß. Anders ausgedrückt: solche

Interviews bringen *verbale Selbstbeurteilungen* zutage, und diese können ziemlich verzerrt sein. - Wir können daraus allgemein schließen: es ist immer zu prüfen, ob in einem Indikator nicht primär ein verstecktes anderes Merkmal gemessen wird.

Die Eigenschaft eines Indikators, jenes theoretische Merkmal zu messen, das er zu messen vorgibt, nennt man in der Statistik auch *Konstruktvalidität*. Man führt oft 'Validitätstests' für Indikatoren durch. Hier ist jedoch Vorsicht geboten: theoretische Merkmale lassen sich nie direkt messen; somit gibt es auch keine Möglichkeit, Konstruktvalidität direkt festzustellen. Um die Korrelation eines empirischen Indikators E mit einem theoretischen Merkmal T zu messen, führt man in der Praxis für T einen anderen empirischen Indikator E* ein und mißt die Korrelation zwischen E und E*. Oder man mißt, wie stark eine ganze Gruppe von Indikatoren, die man alle als Indikatoren des gemeinsamen theoretischen Merkmals T vermutet, untereinander korrelieren. Man verschiebt dadurch die Frage, ob ein Indikator ein theoretisches Merkmal gut mißt, auf die Frage, ob ein ganzes Bündel korrelierender Indikatoren ein theoretisches Merkmal gut mißt. Diese Frageverschiebung ist sicher ein Fortschritt: die Intrakorrelation eines empirischen Indikatorbündels läßt vermuten, daß sie ein gemeinsames zugrundeliegendes theoretisches Merkmal erfassen - aber welches? Ob es jenes ist, das der Forscher im Sinn hat, sprachlich benannt hat und ausgewertet, ist letztlich eine Frage der Anwendung der gesamten Theorie, in dem das theoretische Merkmal eine Rolle spielt.

4. 7 *Theorienbewertung, Theorienvergleich, und Theorienfortschritt*

Unter einer Theorieversion verstehen Kern plus Peripherie, also alle Bestandteile einer Theorie, die nötig sind, um aus der Theorie gehaltvolle empirische Konsequenzen zu gewinnen. Eine Theorie kann im Laufe ihrer Entwicklung verschiedene Versionen durchlaufen, sie kann in ihrer Peripherie angereichert, modifiziert, eingeschränkt werden, usw. Solange nur die Peripherie geändert wird, bleibt es dieselbe "Theorie". Änderung des Theoriekerns führt dagegen zu einer neuen Theorie. Man kann die theoriendynamische Entwicklung also schematisch so darstellen

"Normal
-wissenschaft"

"Revolution"

K₁ K₁

K₂

K₂

V₁ V₂V₁ V₂

Theorie 1

Theorie 2

wobei in einer Disziplin auch häufig mehrere solche Theorien nebeneinander existieren.

Theorien werden anhand ihrer empirischen Konsequenzen überprüft. Dabei erinnern wir uns an das drei-Ebenen-Modell von Kap. 1 : Theorien werden anhand der aus ihnen folgenden empirischen Gesetze überprüft, welche ihrerseits wiederum gemäß den in Kap. 3 besprochenen Methoden anhand von Beobachtungssätzen überprüft werden. Das setzen wir jetzt als unproblematisch voraus und gehen von einer Menge bereits akzeptierter empirischer Gesetzesmäßigkeiten bzw. empirischer Regelmäßigkeiten aus, welche die Theorien zu erklären versuchen. Eine empirische Gesetzesmäßigkeit bzw. Regelmäßigkeit nennen wir auch ein *empirisches Phänomen*.

Wir entwickeln die Begriffe der Bewährung und Schwächung zuerst für Theorieversionen und dann, darauf aufbauend, für Theoriekerne.

1) Eine Theorieversion ist umso mehr bewährt, je mehr empirische Überprüfungsversuche sie erfolgreich bestanden hat (d.h. je mehr empirische Phänomene sie korrekt voraussagt bzw. erklärt).

2) Eine Theorieversion ist streng falsifiziert, wenn mindestens eine ihrer deduktiven empirischen Konsequenzen durch Beobachtungssätze widerlegt wurde.

Eine Theorieversion ist probabilistisch geschwächt, wenn mindestens eine ihrer probabilistischen empirischen Konsequenzen durch Beobachtungssätze widerlegt wurde.

Deduktive Theorieversionen sind streng falsifizierbar; das Holismusproblem besteht darin, daß dadurch nicht bestimmt ist, welche bestandteile der Theorie man preisgibt; im Regelfall werden es Peripheriebestandteile sein. Ein Theoriekern allein ist dagegen nie definitiv falsifiziert; denn man kann durch ad-hoc Hypothesen, z.B. Bereichseinschränkungen oder Ausnahmeklauseln, den Theoriekern immer vor Falsifikation schützen. Daher gibt es kein "experimentum crucis", d.h. kein entscheidendes Experiment für oder gegen einen Theoriekern (Duhem) - zumindest dann nicht, wenn man nicht schon einige theoretische Annahmen als gegeben und unproblematisch voraussetzt.

Unter einer *ad hoc Hypothese* verstehen wir im folgenden einen peripheren Bestandteil einer Theorieversion, welcher nur dazu eingeführt wurde, um einen Widerspruch einer früheren Version der Theorie mit einem empirischen Phänomen zu beseitigen. Solche ad-hoc Hypothese haben wir bereits kennengelernt, z.B. das Postulat eines unbekanntem "störenden" Planeten bei Abweichung der Beobachtung von der vorausgesagten Planetenbahn, oder das Postulat einer zusätzlichen kognitiven Schwierigkeitsvariable bei Abweichung von Testergebnissen von Piagets Stufentheorie. Beachte, eine Hypothese nennen wir nur solange "ad hoc", solange nicht neue empirische Konsequenzen gefunden werden, die sich aus der ad-hoc Hypothese (im Kontext der Theorie) ergeben.

Wir beurteilen die Bewährung oder Schwächung eines Theoriekerns die Geschichte seiner zugehörigen Theorieversionen sowie über seine aktuelle Theorieversion.

- (3) Ein Theoriekern ist umso mehr bewährt,
- je bewährter seine aktuelle Theorieversion ist.
 - je weniger ad-hoc Hypothesen seine aktuelle Theorieversion enthält
 - je mehr empirischen Gehalt seine aktuelle Theorieversion hat.
- (4) Ein Theoriekern ist umso mehr geschwächt,
- je mehr vergangene Theorieversionen bereits verworfen wurden
 - je mehr ad hoc Hypothesen seine aktuelle Theorieversion enthält
 - je weniger empirischen Gehalt seine aktuelle Theorieversion enthält.

Wir können auch *intertheoretische* Vergleichsbewertungen vornehmen; vorausgesetzt ist nur, das es eine beiden Theorien gemeinsame empirische Grundsprache gibt, in der die empirischen Konsequenzen ermittelt und verglichen werden.

Unter einem *Gegenbeispiel* zu einer aktuellen Theorieversion verstehen wir ein empirisches Phänomen, das entweder der Theorieversion direkt widerspricht, oder aber eines, das einer früheren Theorieversion direkt widersprach und in der gegebenen Theorieversion durch eine ad hoc Hypothese im oben erläuterten Sinn vom Anwendungsbereich der Theorie ausgeschlossen wurde.

Wir schreiben eine Theorieversion allgemein als K_iV_j (Version j des Kerns i). Ein einfaches intertheoretisches Vergleichskriterium, das auf die Poppersche Idee der "Wahrheitsnähe" zurückgeht, ist folgendes:

- (5) Theorieversion K_1V_i ist Theorieversion K_2V_j rational vorzuziehen (bzw., der Übergang von K_1V_i zu K_2V_j ist "Theorienfortschritt"), wenn

entweder K_1V_i erklärt mehr empirische Phänomene als K_2V_j und hat nicht mehr Gegenbeispiele als K_2V_j

oder K_1V_i erklärt gleichviel empirische Phänomene wie K_2V_j , hat jedoch weniger Gegenbeispiele

Das "mehr", "gleichviel", "weniger" ist dabei im mengentheoretischen Sinn der Mengeninklusion zu verstehen, nicht im numerischen Sinn des Abzählens, das wäre willkürlich. Obiges Kriterium liefert nur eine *partielle Ordnung*, d.h. läßt es in vielen Fällen unentschieden, welche der beiden Theorieversionen vorzuziehen ist. Man kann natürlich noch weitere Vergleichsparameter heranziehen, wie z.B. der *prognostische Überschußgehalt* - das ist jeder Teil des empirischen Gehalts, der noch nicht getestet wurde; oder die *Einfachheit* der Theorie - dies wird dann sinnvoll, wenn sich aufgrund obigen Kriteriums keine Entscheidung ergibt.

Die Frage, ob eine Theorie bzw. ein Theoriekern rational zu akzeptieren, beizubehalten, oder aufzugeben ist, hängt - wie schon in Kap. 1 ausgeführt - nicht nur vom Stand der empirischen Überprüfung, sondern auch vom Stand der Alternativtheorien ab.

(6) Ein Theoriekern ist rational zurückzuweisen, wenn er stark geschwächt ist Im Sinn von (4), und zu seiner aktuellen Version eine wesentlich bessere Theorieversion (im Sinn von 5) mit anderem Theoriekern vorliegt.

4.8 Die Bedeutung der empirisch-theoretisch-Unterscheidung für das Programm einer objektiven empirischen Wissenschaft

Wir wollen uns die Wichtigkeit der Unterscheidung zwischen Beobachtungsbegriffen und theoretischen Begriffen für das empiristische Programm bzw. das Programm einer objektiv-rationalen Wissenschaft überhaupt klarmachen. Unser Begriff des Beobachtungssatzes entspricht ja keineswegs dem üblichen wissenschaftlichen Begriff des Datums. Z.B. sind in der wissenschaftlichen Sprechweise Sätze wie

(1) Die Masse des Elektrons beträgt $9.10 \cdot 10^{-28}$ g

(2) 50% aller 6-jährigen Kinder befinden sich auf der präoperationellen Stufe des Denkens (gemäß Piaget)

'Daten'. Theoretische Hypothesen können durch derartige 'Daten' scheinbar äußerst simpel überprüft werden - physikalisch-chemische Gesetze durch 'Daten' vom Typ

(1); oder kognitionspsychologische Gesetzhypothesen durch 'Daten' vom Typ (2). Tatsächlich sind (1) und (2) aber *theoretische Singulärsätze*, wogegen die tatsächlichen Beobachtungssätze im Fall (1) Zeigerausschläge und empirische Meßanordnungen, im Fall (2) Resultate von Piagetschen Leistungstests sind (!). Wie aber gelangt man von diesen wirklichen Beobachtungssätzen zu jenen theoretischen Singulärsätzen, die dem üblichen Begriff des wissenschaftlichen Datums entsprechen? Die Antwort lautet: durch einen ganzen *Komplex von stillschweigend vorausgesetzten theoretischen Hypothesen und Zuordnungsgesetzen*. Darüber lassen uns nicht nur die meisten fachwissenschaftlichen, sondern auch (leider) die meisten wissenschaftstheoretischen Lehrbücher im Unklaren: sie unterlassen häufig die *volle Rekonstruktion* des Bezug von Theorien zu Beobachtungssätzen (weil dieser äußerst komplex ist) und belassen es bei obigen 'Daten'. Dann ist man freilich gezwungen, zu sagen, daß derartige 'Daten' selber von theoretischen Annahmen abhängig sind (so z.B. Popper, ebenso Weingartner 1978). Viele Personen werden an dieser Stelle aber sofort weiterfragen: von welchen seltsamen, scheinbar der nicht weiter durchdringbaren Kompetenz von Spezialisten überlassenen theoretischen Kenntnissen und Voraussetzungen hängen solche 'Daten' denn ab? Und wenn alle Daten bereits von theoretischen Voraussetzungen abhängen, dann muß doch jede Überprüfung einer Theorie durch Daten *zirkulär* sein? Dann ist doch der Schluß naheliegend, daß die wissenschaftlichen 'Daten' und somit die ganze Wissenschaft etwas ziemlich Subjektives ist. Tatsächlich haben Kuhn (1976), Feyerabend (1976) und andere Autoren diesen Schluß gezogen. Alle Beispiele von 'Daten', die diese Autoren jedoch zur Stützung ihrer Thesen von der 'Relativität' von wissenschaftlichen Daten heranziehen, sind Beispiele von theoretischen Singulärsätzen, nicht von wirklichen Beobachtungssätzen. Z.B. argumentierte Kuhn, Ptolemäus und Kepler hätten etwas anderes am Himmel gesehen, weil Ptolemäus dort einen Planeten sah, wo Kepler einen Fixstern sah (die Sonne). Man sieht jedoch weder einen Fixstern noch einen Planeten, sondern höchstens einen *leuchtenden Fleck* am Himmel. Ob dieser leuchtende Fleck ein Planet ist, d.h. kreisförmig um einen Zentralkörper wandert, oder ein Fixstern ist, ist eine theoretische, nichtbeobachtbare Frage. - Man sieht jedenfalls: *der Subjektivitäts- oder Relativitätsvorwurf gegen die Wissenschaft läßt sich nur entkräften, wenn man sich der Mühe unterzieht, den Überprüfungszusammenhang tatsächlich bis zur Ebene der elementaren Beobachtungssätze hin zu rekonstruieren.*

Erwähnt sei noch, daß auch die bekannte 'Inkommensurabilitätsthese' von Kuhn - derzufolge verschiedene Theorien gar nicht rational vergleichbar wären, weil sie ganz unterschiedliche Begriffe haben - auf der Voraussetzung beruht, daß man in Theorien nicht zwischen empirischen und theoretischen Anteilen unterscheiden kann, weil alle 'Daten' theoriegeladen sind. Wenn man - wie wir das hier tun - in Theorien eine

empirische Sprache und ihren empirischen Gehalt abgrenzen kann, so fällt diese Kuhnsche These (in ihrer radikalen Version) in sich zusammen, weil dann unterschiedliche Theorien eben über ihren jeweiligen empirischen Gehalt, formuliert in der ihnen gemeinsamen theorieneutralen Beobachtungssprache, vergleichbar sind.

4.9 Erkenntnistheoretische Interpretation von Theorien: Instrumentalismus und Realismus

Eine wissenschaftstheoretisch vieldiskutierte Frage war: welche Erkenntnisfunktion hat eigentlich jener Teil von Theorien, der über ihren empirischen Gehalt hinausgeht? Zumeist bezwecken theoretische Begriffe, die hinter den empirischen Phänomenen verborgenen und unbeobachtbaren Entitäten zu erfassen und mit den theoretischen Gesetzen deren Mechanismen bzw. Kräfte abzubilden (realistische Position), manchmal bezwecken sie aber nur, die empirischen Phänomene möglichst einfach und nützlich zusammenzufassen (instrumentalistische Position). Die Frage, ob die instrumentalistische oder die realistische Position die angemessenere ist, hängt selber zweifellos vom Grad der empirischen Bewährung der Theorie (s.u.) ab: bei einer bewährten Theorie wie der Atomtheorie scheut man sich nicht, den Atomen reale Existenz zuzusprechen; bei einer wackeligen Theorie wie der Quarktheorie oder gar der Jungschen Theorie des 'kollektiven Unbewußten' wird man schon eher fragen, ob "Quark" oder "kollektives Unbewußtes" tatsächlich real existierende Entitäten bezeichnen.

Aus realistischer Perspektive ist mit einer Theorie immer auch ein theoretisches Modell verbunden, welches ein Bild der unbeobachtbaren Entitäten liefert - z.B. die Modellvorstellung von Atomen und Molekülen, oder die psychologische Modellvorstellung von "Trieben" (Dampfkessel-Ventil-Prinzip). Logisch gesprochen drückt sich die Modellvorstellung in theoretischen Existenzsätzen oder theoretischen Existenzallsätzen aus: es gibt die und die kleinsten Materiebestandteile, und alle Substanzen sind aus ihnen aufgebaut; es gibt die und die elementaren Kraftwirkungen zwischen Materiebestandteilen, und alle Bewegungsveränderungen lassen sich darauf zurückführen; es gibt die und die Basistriebe des Menschen, und alle seine emotionellen Bestrebungen lassen sich aus Mischungen dieser Basistriebe erklären; usw.

Analog zur Unterscheidung zwischen empirischen Gesetzen und Theorien ist die aus der Psychologie bekannte Unterscheidung zwischen sogenannten black-box-Hypothesen, die bloß empirische Phänomene miteinander korrelieren, und theoretischen Hypothesen, die theoretische Gesetze und entsprechende Modelle enthalten (Mario Bunge nennt sie auch "phänomenologische" versus "repräsentationale" Hypothesen; vgl. Weingartner 1978, S. 66). Der Behaviourismus

in der Psychologie (jedenfalls in seiner früheren extremen Form) ist eine typische Form einer reinen black-box-Position: er beschränkt sich darauf, Gesetze zwischen den Reizen (stimulus S), die auf einen Menschen (bzw. Organismus O) wirken, mit den darauf folgenden Reaktionen (response R) zu studieren, weil nur diese beobachtbar sind, während über die (unbeobachtbare) innere mentale bzw. psychologische Struktur des menschlichen Fühlens und Denkens keine Aussagen gemacht werden sollen, d.h. kein theoretisches Modell davon gebildet werden soll. - Jene Vertreter der behaviouristischen Psychologie, die sich dabei auf die Naturwissenschaften beriefen, hatten letztere ziemlich mißverstanden, weil diese alles andere als empirisch - vielmehr hochtheoretisch - ist; kein Wunder, daß man die (zumindest streng) behaviouristische Position in der Psychologie bald aufgab, weil sie zu wenig Ertrag zuließ. Behaviorismus ist nichts ist eine Form des "naiven" oder reduktionistischen Empirismus. Unser minimaler Empirismus läßt theoretische Modelle, die das Beobachtbare inhaltlich übersteigen, zu, er fordert lediglich, daß sich diese theoretische Modelle empirische Konsequenzen besitzen müssen, über die sie sich zu bewähren haben oder geschwächt/falsifiziert werden.

4. 10 Letztes Beispiel: Die Adorno-Milgram-Theorie des autoritären Charakters

Wir rekonstruieren abschließend ein Beispiel einer empirischen Überprüfung einer sozialwissenschaftlichen Theorie: die Überprüfung der Theorie der autoritären Gesellschaft durch das Milgram-Experiment. Die soziologische Kernthese dieser Theorie, die unter anderem von Adorno in "The Authoritarian Personality" vertreten wurde, lautet:

K: Unsere Gesellschaft ist in ihrem Innern autoritär strukturiert, das heißt, die Individuen haben die autoritäre Ideologie in Form eines stark autoritätsorientierten Charakters verinnerlicht.

Adorno, aber auch Milgram, wollten mit dieser These voralledem erklären, wie es zum Faschismus kommen konnte. Die These diente Marcuse aber auch dazu, das relative Ausbleiben von Revolten und manifester Gewalt durch den Staat in der spätkapitalistische Gesellschaft zu erklären - eine Gesellschaft, in der Autorität verinnerlicht ist, hat manifeste Gewalt nicht mehr nötig.

Die Idee von Milgram, diesen verinnerlichten Autoritarismus zu prüfen, war folgende: solche Menschen müßten gegenüber von ihnen anerkannten Autoritäten eine starke Tendenz zu bedingungsloser Unterwerfung aufweisen. Formulieren wir dies als Spezialgesetz (es spezifiziert die Bedeutung, ist aber noch kein volles Zuordnungsgesetz, weil es noch keine Verbindung mit beobachtbaren Begriffen

erzeugt - aber sozusagen diesen näherkommt):

S1: Menschen mit stark autoritätsorientiertem Charakter tendieren stark dazu, gegenüber von ihnen anerkannten Autoritäten sich bedingungslos zu unterwerfen.

Die weitere Idee von Milgram, die ihn zu seinem Experiment führte, war nun folgende: man nehme eine gesellschaftlich anerkannte Autorität, z.B. einen wissenschaftlichen Experten. Dieser Wissenschaftler W soll die Vpn (Versuchsperson) zu einer Handlung auffordern, die in seinen Kompetenzbereich fällt - z.B. an einem wissenschaftlich gutgeheißenen Experiment mitzuarbeiten. Die Handlung soll jedoch so geartet sein, daß sie den gesellschaftlich allgemein anerkannten Moralvorstellungen stark widerspricht - z.B. wie bei Milgram: die Vpn wird von W aufgefordert, in einem wissenschaftlich gutgeheißenem Experiment eine andere Person, die bzgl. ihrer Lernfähigkeit trainiert werden soll, durch Elektroschocks negativ zu konditionieren (wobei diese scheinbar gequälte Person in Wirklichkeit durch einen Schauspieler gespielt wurde). Die Vpn sieht die Person über einen Videoschirm, oder hört sie zumindest; glaubt wirklich daß die Person Schmerzen leidet. Sie steht im Konflikt: soll ich meiner moralischen Vorstellung nach handeln, oder soll ich mich an der Autorität von W orientieren, welcher versichert, das Experiment sei trotz der Schmerzzufügung ethisch vertretbar und äußerst wichtig. Dabei hat die Vpn die Freiheit, das Experiment abzubrechen; sie wird von W nicht zum Weiterführen gezwungen, sondern nur verbal dazu zu 'überreden' versucht.- Personen, die in einer solchen Situation, gegen ihre ansonsten akzeptierten Moralvorstellungen, sich der Autorität freiwillig unterwerfen, sind sicher Belege für eine 'Tendenz zur bedingungslosen Unterwerfung'. Somit kommen wir nun zu einem weiteren Spezialgesetz, das uns der Ebene der Beobachtung noch näherbringt:

S2: Menschen M, die stark zur bedingungslosen Unterwerfung gegenüber anerkannten Autoritäten tendieren, und nur solche, werden in folgendem Experimenttyp wie folgt reagieren: Wenn eine von M anerkannte Autorität A M zu einer Handlung H mit Nachdruck auffordert, aber nicht zwingt, wobei H in den Kompetenzbereich (=Bereich der Autoritätskraft) von A fällt, jedoch mit den ansonsten moralisch akzeptierten Normen von M in Konflikt steht, so wird M mit hoher Wahrscheinlichkeit dennoch H tun.

Nun benötigen wir nur noch einige offensichtlichen Indikatorgesetze, um direkt zum Milgram-Experiment zu kommen:

I1: Beispiel für anerkannte Autorität: Wissenschaftler

I2. Beispiel für Handlung H, die in den wissenschaftlichen Kompetenzbereich fällt, aber den ansonsten akzeptierten Normen der VPn widerspricht: Durchführung eines wissenschaftlich gutgeheißenes Experiments, worin einer Person Schmerzen zugefügt werden.

K+S1+S2+I1+I2 zusammen ergeben die Prognose für das Milgram- Experiment, daß ein hoher %-Satz der Menschen tatsächlich das im Milgram-Experiment vorgetäuschte "wissenschaftliche Experiment" durchführen wird, obwohl dabei eine Person (scheinbar) gequält wird.. Tatsächlich erbrachte das Milgram-Experiment dieses Ergebnis. Das ist als empirische Bewährung des gesamten Theoriekomplexes K+S1+S2+I1+I2 anzusehen.

Allerdings müßte man nun weitere Tests mit variierenden Indikatorbeziehungen anstellen, um tatsächlich den Theoriekern insgesamt als bewährt ansehen zu können. Ansonsten könnte es sich ja nur um ein Spezialphänomen der wissenschaftlichen Expertokratie handeln, aber nicht auf Autoritäten oder Handlungen beliebiger Art übertragbar sein. Andere solche Indikatoren bzw. Überprüfungsbeispiele wären: Ein Offizier des Bundesheers sagt: diese Aktion zerstört zwar große Teile des Wienerwaldes, ist aber für die österreichische Sicherheit unerlässlich - werden dann die meisten Bh.-ler mitmachen? Oder: ein Priester sagt: zu Vertreibung des Bösen aus diesem Menschen ist es nötig, ihn in der und der Art zu foltern, obwohl es ihn quält - werden dann Gläubige mitmachen? - Falls ein anderer solcher Test ein negatives Ergebnis bringen würde, wäre - gemäß dem Holismus der Falsifikation - folgendes möglich: (a) die leitende Person ist keine anerkannte Autorität (I1 falsch); (b) die Handlung fällt nicht in den Bereich, für den die Person als kompetent gehalten wird, oder sie widerspricht nicht wirklich den ansonsten anerkannten Moralvorstellungen (I2 falsch); (c) man zweifelt S2 oder S1 an, am ehesten hier noch S1: es könnte ja sein, daß die Unterwerfungstendenz nur gegenüber bestimmten Autoritäten, z.B. nur gegenüber scheinbar 'objektiven Experten', nicht aber gegenüber Politikern etc., auftritt; (d) schließlich könnte man die Kernthese selber anzweifeln. Falls hingegen der ganze Komplex K+S1+S2+I*1+I*2 mit den neuen Indikatorbeziehungen I*1 und I*2 (für Bh.-Befehlshaber oder Priester) ebenfalls empirisch bewährt, hat man guten Grund, die Kernthese K als bewährt anzusehen, und somit auch als zutreffende Erklärung z.B. für den Faschismus anzusehen.

4.11 Literaturhinweise:

Klassiker zu wissenschaftlichen Theorien - und einfach zu lesen - sind wieder die Werke von Hempel und Carnap in Kap. 1. Ebenfalls findet sich im Band II der Stegmüller-Reihe die Entwicklung des Theoriebegriffs diskutiert. Band II, 3.

Teilband (Die Entwicklung des neuen Strukturalismus seit 1973), erschienen 1985, gibt den erwähnten Überblick über die strukturalistische Theorienauffassung. Zu den Werken von Kuhn und Feyerabend s. wieder Kap. 1. Ferner: I. Lakatos, A. Musgrave (Hg., 1974), Kritik und Erkenntnisfortschritt, Vieweg, Braunschweig - darin der Aufsatz von Lakatos "Falsifikation und die Methodologie wissenschaftlicher Forschungsprogramme. Zu naturwissenschaftlichen Theorien: Duhem, P. (1978, Nachdr. d. Aufl. von 1908): Ziel und Struktur der physikalischen Theorien, F. Meiner, Hamburg. Sneed, J. D. (1971): The Logical Structure of Mathematical Physics, Reidel, Dordrecht. - Zu human- und sozialwissenschaftlichen Theorien: Gadamer, V. (1984): Theorie und Erfahrung in der psychologischen Forschung, J.C.B. Mohr, Tübingen. Die zitierten Aufsätze: Schurz, G. (1985): "Denken, Sprache und Erziehung: Die aktuelle Piaget-Kontroverse", Zeitschrift für Semiotik 7/4, 335-366; (1987): "Der neue Strukturalismus", Conceptus 21/52; (1989): "Paradoxical Consequences of Balzer's and Gähde's Criterion of Theoreticity", in: Erkenntnis 1990.