

Forschungsmethoden der Psychologie

Hochschullehrbücher

Band 2

Wilhelm Kempf

Forschungsmethoden der Psychologie

**Zwischen naturwissenschaftlichem Experiment und
sozialwissenschaftlicher Hermeneutik**

Band 1: Theorie und Empirie

verlag irena regener berlin

Das vorliegende Buch entstand im Rahmen des Projektes "Modulare Methodenausbildung im Fach Psychologie" und wurde aus Mitteln des Ausschusses für Lehre und Weiterbildung an der Universität Konstanz (ALW) gefördert.

3. Auflage
© 2009 by verlag irena regener
ostseestr. 109, d – 10409 berlin
<http://www.regener-online.de>

Alle Rechte vorbehalten, auch die der fotomechanischen Wiedergabe und der Speicherung in elektronischen Medien.

Printed in Germany

ISBN 978-3-936014-21-1

Inhalt

Vorwort	9
-------------------	---

Kapitel 1: Psychologie als Erfahrungswissenschaft

1.1	Die Anfänge der wissenschaftlichen Psychologie	13
1.1.1	Die Überwindung der Vermögenspsychologie	13
1.1.2	Die ersten psychologischen Experimente	16
1.2	Wissenschaftliche Psychologie und psychologisches Alltagswissen	26
1.3	Propädeutische Grundlagen	30
1.3.1	Beobachtungssprache und Theoriesprache	31
1.3.2	Terminologische Regeln	37
1.3.3	Exkurs: Der Streit um den Aggressionsbegriff	42
1.3.4	Die logischen Partikel	48
1.3.5	Wahrheit und Wirklichkeit	52
1.4	Aufgabenverständnis, Gegenstandsverständnis und Wissensideal der Psychologie	62

Kapitel 2: Hard Science: Das Erbe der Naturwissenschaften

2.1	Das aristotelische Wissensideal	73
2.1.1	Die axiomatische Begründung der klassischen Testtheorie	74
2.1.2	Anforderungen an Axiomensysteme.	76
2.1.3	Kritik und Neubegründung der klassischen Testtheorie	79
2.1.4	Exkurs: Herleitung der Axiome von Gulliksen	81
2.2	Das galileische Wissensideal	86
2.3	Das deduktiv-nomologische Erklärungsmodell	88

2.4	Messung	94
2.4.1	Qualitative und quantitative Variablen	94
2.4.2	Die Guttman-Skala	97
2.4.3	Skalentypen	100
2.4.4	Allgemeine und angewandte Messtheorie	102
2.4.5	Manifeste und latente Variable	109
2.4.6	Latent-Trait-Modelle	113
2.4.7	Exkurs: Informationsverarbeitungsmodelle der kognitiven Entwicklung	121
2.4.8	Latent-Class-Modelle	127
2.5	Experimentelle Hypothesenprüfung.	141
2.5.1	Grundbegriffe des Experimentierens	143
2.5.1.1	Exkurs: Das Ferienlagerexperiment von Sherif	152
2.5.1.2	Exkurs: Das Milgram-Experiment	154
2.5.2	Voraussetzungen experimenteller Hypothesenprüfung	157
2.5.2.1	Widerspruchsfreiheit	160
2.5.2.2	Kritisierbarkeit	163
2.5.2.3	Exkurs: Die Triebtheorie der Aggression nach Konrad Lorenz	167
2.5.2.4	Kontingenz	170
2.5.2.5	Operationalisierbarkeit.	175
2.6	Das induktiv-statistische Erklärungsmodell.	182
2.6.1	Zufall und Wahrscheinlichkeit.	184
2.6.1.1	Die konstruktive Begründung des Wahrscheinlichkeitsbegriffs	185
2.6.1.2	Exkurs: Das Gesetz der großen Zahlen	190
2.6.2	Statistische Modellbildung	194
2.6.3	Statistische Inferenz	199
2.6.3.1	Die Schätzung unbekannter Parameter	201
2.6.3.2	Das Testen von Hypothesen.	207
2.6.3.3	Zweiseitige und einseitige Fragestellungen	213
2.6.4	Zufallsauswahl und Zufallsexperiment.	217

Kapitel 3:

Soft Science: Die Herausforderung an die Sozialwissenschaften

3.1	Objektseitig und subjektseitig definierte Variablen	227
3.1.1	Exkurs: Materielle und strukturelle Reduktion	228
3.1.2	Subjektseitig definierte Beschreibungsprädikatoren	230
3.1.3	Verhalten, Handlung und Widerfahrnis	234
3.1.4	Ziel und Mittel.	235

3.2	Das intentionale Erklärungsmodell	238
3.2.1	Bedingungsbeziehungen vs. Begründungsbeziehungen	239
3.2.2	Die Hypothesenförmigkeit intentionaler Erklärungen	241
3.2.3	Die empirische Prüfbarkeit intentionaler Erklärungen	244
3.2.4	Exkurs: Der Unkritizierbarkeitsverdacht gegenüber der Psychoanalyse	249
3.3	Das narrative Erklärungsmodell	254
3.3.1	Reflexionstermini	254
3.3.2	Die narrative Erklärung	255
3.3.3	Sozialwissenschaftliche Hermeneutik.	257
3.4	Sinnrationalität	260
3.4.1	Regelgeleitetes Handeln	260
3.4.2	Nonfinale Handlungsorientierungen	261
3.4.3	Exkurs: Die Perspektivität des Gewaltbegriffs.	263
3.4.4	Kontextabhängigkeit	269
3.4.5	Metakommunikative Axiome.	272
3.5	Die systemtheoretische Erweiterung des handlungstheoretischen Modells	281
3.5.1	Der Begriff des Systems.	281
3.5.2.	Selbstreferentialität	282
3.5.3	Kybernetische Regelungsprozesse.	287
3.5.4	Schärfegrade des autonomen Prozesses	289
3.6	Informationsverarbeitungsmodelle	293
3.6.1	Der Begriff der intervenierenden Variablen	293
3.6.2.	Die Signal-Entdeckungs-Theorie	295
3.6.3	Sieglers Theorie der kognitiven Entwicklung im Kindesalter	299
3.6.4	Die statistische Regelmäßigkeit des Entwicklungsverlaufes	305
3.7	Strukturelle und empirische Theorieanteile	311
3.7.1	Der Zirkularitätsverdacht gegenüber dem Gesetz des Effektes	311
3.7.2	Zirkularität, Pseudoempirie und strukturelle Theorieanteile	316
3.7.3	Frustration und Aggression	317
3.7.4	Exkurs: Die Lerntheorie der Aggression	322

3.8	Handlungstheoretische Prozessanalyse.	326
3.8.1	Die Erfahrungsgebundenheit struktureller Theorieanteile	326
3.8.2	Exkurs: Die Dynamik der Konflikteskalation	327
3.8.3	Narrative Erklärung, Eskalationslogik und Begründungsmuster	329
3.9	Qualitative Forschungsmethoden.	335
3.9.1	Interpretativität	337
3.9.2	Text und Textanalyse.	341
	Literatur	348
	Anhang	
	Über den Autor	363
	Statistische Tabellen	364
	Aufgabenlösungen	366
	Sachregister	369

Vorwort

Das vorliegende Buch ist der erste Band eines dreibändigen Lehrbuchs der psychologischen Methodenlehre und entstand im Rahmen eines Projektes zur modularen Methodenausbildung im Fach Psychologie an der Universität Konstanz.

In didaktischer Hinsicht von der Überzeugung geprägt, dass psychologische Forschungsmethoden nicht losgelöst von den Inhalten der Psychologie unterrichtet, geschweige denn verstanden werden können, fasst das Lehrbuchwerk die Psychologie als Wissenschaft von der subjektiven Welt des Menschen, deren Gegenstandsbereich von den physiologischen Ursachen psychischer Phänomene bis hin zu den Prozessen der sozialen Konstruktion von Wirklichkeit reicht. Entsprechend verfolgt es eine wissenschaftstheoretische Konzeption, welche die Methoden der Psychologie im Spannungsfeld zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik verortet und eine gegenstandsangemessene Umsetzung des galileischen Denkens in der Psychologie zum Ziel hat.

Im Zentrum dieses mit *Theorie und Empirie* betitelten ersten Bandes steht daher nicht nur die Vermittlung der Grundbegriffe von Methodologie und Methodik, sondern zugleich auch die Herstellung des Zusammenhanges zwischen den verschiedenen Spezialgebieten der psychologischen Methodenlehre und die Herstellung ihres Zusammenhanges mit den substanzwissenschaftlichen Inhalten der Psychologie. Um dieses Anliegen umzusetzen, werden die behandelten Methodenprobleme nicht nur aus ihrer methodologischen Systematik, sondern auch aus der Psychologie selbst heraus entwickelt, wofür insbesondere auf klassische psychologische Theorien, Experimente und Untersuchungen aus den Gebieten der Psychophysik (Weber'sches Gesetz), der Lerntheorie (klassisches und operantes Konditionieren), der differenziellen Psychologie (Intelligenz- und Einstellungsmessung), der kognitiven Psychologie (Informationsverarbeitungsprozesse) und der Sozialpsychologie (Gruppenprozesse, Aggressions- und Konfliktforschung) zurückgegriffen wird.

Einzelne Kapitel des Buches setzen bereits gewisse Kenntnisse der Wahrscheinlichkeitsrechnung und Statistik voraus. Diese sind entsprechend gekennzeichnet und können beim ersten Durcharbeiten des Buches ohne Verlust der Kontinuität übersprungen werden. Insofern, als das vorliegende Buch eine Einführung in die Methoden der Psychologie darstellt, erwies es sich als unvermeidbar, dass manche der behandelten Methodenfragen eine z. T. vereinfachende Darstellung erfahren bzw. nicht jedes Methodenproblem bis ins letzte Detail diskutiert werden kann. Dass die Korrektheit der Darstellung dadurch nicht eingeschränkt wird, ist der Anspruch des Autors.

Zwei der bereits hier angesprochenen Themenbereiche erfahren in dem mit *Quantität und Qualität* überschriebenen zweiten Band der "Forschungsmethoden" eine vertiefte Behandlung: Zum einen Textinterpretation und Inhaltsanalyse, die dem Wissensideal der qualitativen Sozialforschung verpflichtet sind, und zum anderen die Psychometrie, welche die wohl konsequenteste Umsetzung eines quantifizierenden Wissensideals in der Psychologie darstellt. Den Antagonismus zwischen quantitativen und qualitativen Forschungsmethoden überwindend, will das Buch den Leser in die Lage versetzen, die einschlägige Fachliteratur kritisch zu rezipieren, die Methoden in der eigenen Forschungspraxis anzuwenden und sie in gegenstandsangemessener Weise miteinander zu verbinden. Nicht nur die Anwendungsnähe der Methoden, sondern auch ihre theoretische und (im Falle der Psychometrie) mathematische Grundlegung stehen daher im Zentrum des Buches.

Eine ähnlich integrative Intention verfolgt auch der gemeinsam mit Markus Kiefer herausgegebene dritte Band der "Forschungsmethoden", der den Titel *Natur und Kultur* trägt. In die wichtigsten Verfahren der qualitativen Sozialforschung, der behavioralen experimentellen Psychologie und der Neurowissenschaften einfürend, plädiert das Buch für den gegenstandsbezogenen Einsatz verschiedenster Forschungsmethoden, der die klassischen Grenzen zwischen natur- und kulturwissenschaftlichen Herangehensweisen ebenso aufhebt wie jene zwischen quantitativen und qualitativen Forschungsmethoden.

Insgesamt gibt das Lehrbuchwerk damit nicht nur eine Einführung in das Methodenspektrum der verschiedenen psychologischen Strömungen, sondern entwirft die Konzeption einer integrativen Psychologie, in der die dem Gegenstand der Psychologie inhärente Dialektik von Subjektivität und Objektivität zur Synthese gebracht wird.

Kapitel 1

Psychologie als Erfahrungswissenschaft

1.1

Die Anfänge der wissenschaftlichen Psychologie

1.1.1 Die Überwindung der Vermögenspsychologie

Die Psychologie ist eine relativ junge Wissenschaft, die sich erst um die Mitte des 19. Jahrhunderts aus der Philosophie losgelöst und als eigenständige Erfahrungswissenschaft (= empirische Wissenschaft) zu konstituieren begonnen hat. Bis dahin hatte man sich die Erklärung psychischer Phänomene denkbar leicht gemacht, indem man die im allgemeinen Sprachgebrauch bereits vorhandene Klassifizierung – Denken, Fühlen, Wollen – übernommen und diese drei Verhaltensweisen als Fähigkeiten der Seele aufgefasst hat.

Vermögens-
psychologie

Dass diese sog. Vermögenspsychologie, die bereits von Aristoteles (384–322 v. Chr.) vertreten wurde, keine Erklärung der psychischen Phänomene zu leisten vermag, erscheint uns heute als evident. Offensichtlich handelt es sich dabei um einen logischen Zirkelschluss, der auf einer unzulässigen Verdoppelung der Realität (Holzkamp, 1965) beruht: Klassifikationen von Verhaltensweisen (z.B. Denken) werden vergegenständlicht, und die vergegenständlichten Verhaltensklassifikationen (z.B. Denkvorgänge) werden wiederum als Ursachen des Verhaltens gedeutet (Bergius, 1965; Skinner, 1953; Holt, 1931).

Verdoppelung der
Realität

Gleichwohl wurde die Vermögenspsychologie noch im 18. Jahrhundert von den Aufklärungsphilosophen Christian Wolf (1679–1754) und Nikolaus Tetens (1736–1807) vertreten und sogar von Immanuel Kant (1724–1804) übernommen, der allerdings eine weitere Differenzierung vornahm, indem er das Erkenntnisvermögen in Sinnlichkeit und Verstand einteilte. Erst Herbart (1776–1841) hat sich scharf gegen die Vermögenspsychologie gewandt. Seither gilt sie im Allgemeinen als unwissenschaftlich.

Das bedeutet aber nicht, dass sie definitiv überwunden wäre. Insbesondere im Alltagsleben verwenden wir häufig noch solche Pseudoerklärungen. "Spielt jemand besonders gut Klavier, so

Weiterleben der
Vermögenspsycholo-
gie im Alltagsdiskurs

wollen wir wissen, warum er denn so gut Klavier spielen kann. Sagt nun jemand zu uns, dieser Mensch habe eine ‚musikalische Begabung‘, so neigen wir dazu, diese Redeweise als ‚Erklärung‘ zu akzeptieren. Wir sind im Allgemeinen zufrieden mit der ‚Erkenntnis‘: Er spielt so gut Klavier, weil er musikalisch begabt ist.“ (Werbik, 1974, 49)

Auch in der wissenschaftlichen Psychologie finden nach wie vor Konzepte Verwendung, die – zumindest auf den ersten Blick – wie ein Relikt aus der Vermögenspsychologie anmuten: Das gilt insbesondere für die Konzepte der differentiellen Psychologie, wie z.B. Intelligenz, aber auch z.B. für die – inzwischen als überholt geltende – Annahme eines Aggressionstriebes.

Überwindung der
Vermögens-
psychologie

Dass – und unter welchen methodischen Voraussetzungen – es sich bei diesen Konzepten nicht um eine bloße Verdoppelung der Realität handeln muss, werden wir weiter unten diskutieren, nachdem wir zuerst einen genaueren Blick auf den Begriff der wissenschaftlichen Erklärung – bzw. die verschiedenen Erklärungsbegriffe, welche in der Psychologie nebeneinander bestehen und einander wechselseitig ergänzen – geworfen haben. Hier soll zunächst nur soviel festgehalten werden, dass man in der Abwendung von der Vermögenspsychologie den entscheidenden Fortschritt zu sehen hat, der die erfahrungswissenschaftliche Psychologie erst möglich gemacht hat. Die Bedeutung dieses Fortschrittes liegt darin, "dass man sich nicht mehr mit der einfachen und bequemen Ansicht zufrieden gibt, dass der Mensch die Fähigkeit des Denkens, Wollens, Fühlens usw. besitze, sondern dass man sich bemüht, das *Zustandekommen* der Gedanken, Gefühle, Willensakte usw. genau zu untersuchen. Wie sich die Biologie nicht damit begnügt, den Organismen die Fähigkeit der Bewegung und Fortpflanzung zuzuschreiben und aus diesen Fähigkeiten alles zu erklären, sondern genau erforscht, wie der Bewegungsapparat arbeitet und wie die Organe der Fortpflanzung funktionieren, genauso wenig darf die Psychologie bei der Annahme von ‚Fähigkeiten‘ stehen bleiben". Die psychischen Phänomene müssen vielmehr "in ihrer Entstehung aufgeklärt und bis in ihre gehirnphysiologischen Grundlagen verfolgt werden" (Rohracher, 1963, 61).

Psychologie als
Wissenschaft von der
subjektiven Welt des
Menschen

Dass dies gleichwohl *nicht* bedeutet, dass die Psychologie auf reine Gehirnphysiologie reduziert werden könnte, ergibt sich unmittelbar aus dem Gegenstand der Psychologie, die Rohracher (1963, 102) als "die Wissenschaft von der subjektiven Welt des Menschen" definiert.

Wenn wir uns dieser Definition anschließen, so wird deutlich, dass wir es bei der Psychologie mit einer Wissenschaft zu tun haben, die weit größere und komplexere methodische Schwierigkeiten zu überwinden hat als jede andere Wissenschaft vor ihr. "Will man Psychologie als Wissenschaft betreiben, so muss man alles, was man behauptet, *beweisen*, und man muss überdies die Begriffe, die man verwendet, genau bestimmen" (Rohracher, 1963, 2). Denn solange wir nicht über eine geklärte Wissenschaftssprache (= Terminologie) verfügen, bleibt unklar, was wir mit unseren psychologischen Theorien denn überhaupt behaupten; und es bleibt ebenso unklar, welche Erfahrungstatsachen diese Theorien stützen oder welche sie widerlegen. Zur psychologischen Methodenlehre gehören daher nicht nur die wissenschaftlichen Erklärungsmodelle, derer sich die Psychologie bei ihrer Theoriebildung bedient, sondern auch die Methoden der Geltungsprüfung von Theorien und – zuallererst – die Methoden des Aufbaus einer psychologischen Wissenschaftssprache.

Der grundsätzliche Unterschied zwischen den Forschungsmethoden der Psychologie und der klassischen Naturwissenschaften (Physik, Chemie, Biologie) besteht darin, dass wir dort Sachverhalte untersuchen, die außerhalb des eigenen Erlebens liegen. D.h., dass die Sachverhalte *objektseitig definiert* sind (Hoyningen-Huene, 1989, 43ff). Sie können vom Außenstandpunkt eines neutralen Beobachters her erfasst werden. In der Psychologie trifft dies jedoch nur teilweise zu, z.B. wenn wir die durch einen bestimmten Sinnesreiz (= Stimulus) ausgelöste elektrische Aktivität bestimmter Gehirnzellen untersuchen. In weiten Teilen der Psychologie haben wir es jedoch mit Sachverhalten zu tun, die *subjektseitig definiert* sind und nur vom Innenstandpunkt des erlebenden Subjekts selbst her erfasst werden können. Dies wirft das Problem auf, ob Aussagen über solche Sachverhalte überhaupt anhand objektivierbarer Erfahrungen überprüft werden können – oder ob sich die Psychologie von ihrem Anspruch wieder verabschieden muss, eine empirische Wissenschaft zu sein.

Objektseitig und subjektseitig definierte Sachverhalte

In der Geschichte der Psychologie hat dieses Problem zu einer Reihe von – teilweise bis heute andauernden – Kontroversen geführt, die naturwissenschaftliche vs. geisteswissenschaftliche Psychologie, Erklären vs. Verstehen und/oder quantitative vs. qualitative Forschungsmethoden einander als unversöhnliche Gegensätze gegenübergestellt haben. Für die Entwicklung der Psychologie als Wissenschaft hat sich diese Polarisierung als wenig

Gegenstandsangemessene Umsetzung des galileischen Denkens

hilfreich erwiesen. Während die Mehrheit der naturwissenschaftlich orientierten Psychologen es dem Vorbild der galileischen Physik gleichzutun versuchte und darüber die speziellen Methodenprobleme, welche der Gegenstand der Psychologie mit sich bringt, weitestgehend ignorierte, nahm eine Minderheit von geisteswissenschaftlich orientierten Psychologen diese Probleme zum Anlass, den erfahrungswissenschaftlichen Anspruch der Psychologie schlichtweg abzulehnen – oder salopp gesagt: das Kind mit dem Bade auszuschütten. Die von Lewin (1931) erhobene Forderung nach einer gegenstandsangemessenen Umsetzung der galileischen Denkweise in der Psychologie wurde lange Zeit vernachlässigt.

Wenn die vorliegende Einführung den Versuch unternimmt, diese gegensätzlichen Positionen miteinander zu versöhnen und die Forschungsmethoden der Psychologie im Spannungsfeld zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik zu entwickeln, so werden wir notwendigerweise auch auf methodologische Ergebnisse zurückgreifen müssen, die (noch) nicht zum Standardrepertoire der Psychologieausbildung gehören, sondern benachbarten Disziplinen entstammen: der Logik und der Wissenschaftstheorie, der Mathematik und der Statistik ebenso wie der Biologie und der Soziologie.

1.1.2 Die ersten psychologischen Experimente

Wilhelm Wundt

Als Begründer der modernen Psychologie wird gemeinhin Wilhelm Wundt (1832–1920) genannt, der 1879 in Leipzig das erste Institut für experimentelle Psychologie eingerichtetete. Doch die Einbeziehung erfahrungswissenschaftlicher Methoden in die Psychologie reicht weiter zurück.

David Hume

Schon der englische Philosoph David Hume (1711–1776) vertrat die Auffassung, dass – auch in der Psychologie – nur die Erfahrung zu richtigen Erkenntnissen führen könne. Diese könne sowohl äußere als auch innere Erfahrung sein, wofür Hume die Unterscheidung zwischen Sensation und Reflexion einführte. Bereits für Hume war der Begriff der Seele damit obsolet, denn die Existenz einer Seele kann nicht empirisch bewiesen werden. Durch Erfahrung beweisbar sind nur die Perzeptionen (= die Eindrücke, die wir von der Außenwelt und von unserem bewußten Erleben empfangen) und die Assoziationen (= die Verbindungen

zwischen diesen Eindrücken). So ist das bewußte Erleben für Hume "nichts als ein Bündel oder Zusammen verschiedener Perceptionen, die einander mit unbegreiflicher Schnelligkeit folgen und beständig in Fluß und Bewegung sind" und die durch wechselseitige Beziehungen miteinander zu einer Einheit, dem "Ich", verbunden sind (zit. n. Rohracher, 1963, 4).

Von da an dauerte es noch einmal fast hundert Jahre, bis der Physiologe Ernst Heinrich Weber (1795–1858) und der Physiker Gustav Theodor Fechner (1801–1887) die ersten psychologischen Experimente durchführten und damit der erfahrungswissenschaftlichen Auffassung der Psychologie auch praktisch zum Durchbruch verhalfen.

Die ersten psychologischen Experimente

In seiner Schrift *De tactu* berichtete Weber im Jahre 1834 über Versuche, bei denen er jemandem verschiedene Gewichte auf den Handrücken legte. Dabei stellte er fest, dass nicht jede Erhöhung des Gewichts von seinen Probanden (= Versuchspersonen) auch bemerkt wurde. Damit der Unterschied bemerkt wurde, musste das Gewicht (Vergleichsreiz) umso mehr erhöht werden, je schwerer das vorangegangene (Standardreiz) war:

- So wurde z.B. ein Gewicht von 110g nicht als schwerer empfunden als ein davor aufgelegtes Gewicht von 100g. Auch der Unterschied zwischen einem Vergleichsreiz von 120g und einem Standardreiz von 100g wurde nicht bemerkt.
- Der Unterschied zu einem Standardreiz von 100g wurde erst wahrgenommen, wenn der Vergleichsreiz auf etwa 133g erhöht wurde, also um 33g schwerer war.

Diesen eben merklichen Unterschied zwischen zwei Stimuli bezeichnet man als (absolute) Unterschiedsschwelle und sagt, die Unterschiedsschwelle bei einem Standardreiz von 100g liege bei 33g. Weber fand aber noch mehr:

Absolute Unterschiedsschwelle

- Betrug der Standardreiz 200g, so wurde eine Erhöhung um 33g von den Probanden noch nicht bemerkt, sondern erst eine Gewichtszunahme um 66g. Bei einem Standardreiz von 500g war eine Erhöhung um 160g erforderlich, damit der Unterschied wahrgenommen wurde, und bei einem Standardreiz von 1000g lag die Unterschiedsschwelle bei 320g.
- Ein einfacher Vergleich der Zahlenverhältnisse zwischen Unterschiedsschwelle und Standardreiz zeigt, dass die Gewichte immer um ca. ein Drittel erhöht werden mussten, damit der

Unterschied wahrgenommen wurde (vgl. Tab. 1.1.1).

Tab. 1.1.1: Konstanz der relativen Unterschiedsschwelle beim Auflegen von Gewichten auf den Handrücken

ΔS	S	$\Delta S/S$
33	100	0.33
66	200	0.33
160	500	0.32
320	1000	0.32

$$\Delta S/S = c$$

Relative Unterschiedsschwelle

Dieses konstante Verhältnis (c), das als Quotient zwischen dem eben merklichen Unterschied (ΔS = absolute Unterschiedsschwelle) und dem Standardreiz (S) definiert ist, bezeichnet man als relative Unterschiedsschwelle. Ähnliche Ergebnisse ergaben sich auch auf anderen Sinnesgebieten. Stets musste der Reiz in einem konstanten Verhältnis verstärkt werden, damit der Unterschied wahrgenommen wurde.

Das Weber'sche Gesetz

Mit der (ungefähren) Konstanz der relativen Unterschiedsschwelle (im Bereich mittlerer Reizintensitäten) hatte Weber eine allgemeine Gesetzmäßigkeit gefunden, die später von Fechner mit unendlicher Geduld (z.B. durch 24.576 Gewichtsvergleiche) nachgeprüft und zu Ehren Webers als Weber'sches Gesetz bezeichnet wurde.

Das Weber'sche Gesetz

Der eben merkbare Unterschied zweier Stimuli (ΔS) ist proportional zum Standardreiz (S):

$$\Delta S = c \times S$$

Die Größe der relativen Unterschiedsschwelle (c) hängt dabei sowohl von der Sinnesmodalität als auch von der Art der Versuchsanordnung ab:

- Bei Gewichten, die auf die Hand aufgesetzt werden, beträgt sie etwa $1/2$ bis $1/3$. Beim Gewichtheben ist die relative Unterschiedsschwelle dagegen viel feiner, nämlich $1/50$ bis $1/70$.
- Bei der Lautstärke von Tönen beträgt sie etwa $1/3$, bei der Helligkeit von Lichtreizen etwa $1/100$ bis $1/120$.
- Beim Längenvergleich zweier nacheinander gezeigten Linien beträgt die relative Unterschiedsschwelle $1/50$ bis $1/60$.

Die außerordentliche biologische Bedeutung des Weber'schen Gesetzes erläutert Rohracher (1963, 107f), indem er dessen Auswirkungen auf die Wahrnehmung vergegenwärtigt:

Die biologische Bedeutung des Weber'schen Gesetzes

Das Weber'sche Gesetz

"besagt ja zunächst nichts anderes, als dass die Empfindlichkeit unserer Sinnesorgane für Stärkeunterschiede um so gröber wird, je stärker die Reize sind, um so feiner, je schwächer die Reize werden. Das war aus dem Alltag längst bekannt; die Sterne sehen wir bei Tag deshalb nicht, weil der Reizzuwachs, den sie zum Sonnenlicht darstellen, zu klein ist, um noch eine merkbare Empfindung auszulösen; andererseits hören wir in der Stille der Nacht Geräusche, die im Lärm des Tages unbemerkt bleiben ... Das Geheimnisvolle des Weber'schen Gesetzes liegt aber darin, dass bei einem starken Reiz ungefähr der gleiche Bruchteil ... als Zuwachs eine Stärkeempfindung erzeugt, der auch bei einem schwachen Reiz eine solche hervorruft. Dadurch wird erreicht, dass die Unterschiede zwischen hell und dunkel, laut und leise in unserer Welt in einem mittleren Bereich konstant bleiben: die Gegenstände unserer Umgebung, deren Formen wir am Hell und Dunkel und an ihren Schattenwirkungen erkennen, behalten diese Unterschiede im hellen Sonnenlicht und in der Dämmerung in annähernd gleichem Maße bei. Der Unterschied zwischen dem Weiß des Papiers und dem Schwarz der Buchstaben erscheint uns im hellen Mittagslicht ungefähr gleich groß wie im Dämmerlicht des Morgens; zum mindesten können wir zu beiden Zeiten mühelos lesen. In Wirklichkeit ist aber, wie Hering gezeigt hat, das Schwarz der Buchstaben um die Mittagszeit objektiv ungefähr dreimal heller als das Weiße des Papiers am Morgen. Wäre ... die relative Unterschiedsempfindlichkeit nicht konstant, so wären das Weiß und das Schwarz zueinander nicht mehr im gleichen Verhältnis und wir könnten nur bei einer bestimmten Helligkeit lesen, die Gegenstände unserer Umgebung nur zu bestimmten Tageszeiten genau unterscheiden und wiedererkennen".

Die wissenschaftshistorische Bedeutung des Weber'schen Gesetzes für die Entwicklung der Psychologie zur empirischen Wissenschaft kann ebenfalls nicht hoch genug eingeschätzt werden: Mit seiner Entdeckung war es erstmals gelungen, eine Art Naturgesetz zu finden, welches die subjektive Welt des Menschen nicht einfach als Abbild der objektiven Wirklichkeit mit dieser gleichsetzt, aber gleichwohl einen Wirkungszusammenhang angibt, der die Entstehung dieser subjektiven Welt (d.h. hier: die subjektiv definierte Wahrnehmung von Gleichheit oder Unterschiedlichkeit zweier Stimuli) aus materiellen Gegebenheiten der objektiven Wirklichkeit (d.h. hier: aus den objektseitig definierten Unterschieden in der Reizintensität) zu erklären vermag.

Die wissenschaftshistorische Bedeutung des Weber'schen Gesetzes

Man kann sich leicht vorstellen, welche Signalwirkung die Entdeckung einer solchen Gesetzmäßigkeit, die sich zudem noch ma-

thematisch formulieren ließ, für die Entstehung der experimentellen Psychologie entfalten musste. So führte Webers Entdeckung nicht nur zur Entstehung der von Fechner begründeten Psychophysik, sondern auch in anderen Teilgebieten der Psychologie hielt das Experiment Einzug.

Selbstbeobachtung
vs.
Operationalisierung
psychischer
Phänomene

Für die Experimentalpsychologen des 19. Jahrhunderts stellte dabei die Selbstbeobachtung noch die wichtigste Methode der Psychologie dar, d.h. die Beschreibung des eigenen bewußten Erlebens. Die Versuchspersonen (Vpn) waren meist eigens darin ausgebildet und in der Regel wissenschaftliche Mitarbeiter und Studenten. Als Zugang zur subjektiven Welt des Menschen scheint die Selbstbeobachtung in der Tat auch *die* Methode der Wahl zu sein. Gleichwohl wirft die Selbstbeobachtung eine Reihe von Methodenproblemen auf, die man entweder zu lösen oder zu umgehen versuchen kann. Zu diesen Problemen gehören grundsätzliche Fragen, wie z.B. die nicht auszuschließende Möglichkeit der Veränderung des subjektiven Erlebens durch den Prozeß der Selbstbeobachtung ebenso wie eher technische Fragen, wie z.B. die Schwierigkeiten der eindeutigen Klärung subjektseitig definierter Termini und/oder der Messung subjektseitig definierter Größen. Solche Methodenprobleme veranlassten schon die Mitarbeiter Wundts um 1885 dazu, Apparate in die psychologische Methodik einzuführen, welche es ermöglichen sollten, subjektseitig definierte Sachverhalte (z.B. emotionale Reaktionen) durch objektseitig definierte Indikatoren (z.B. Veränderungen der Puls- und Atmungsfrequenz) zu operationalisieren und damit von einem unbeteiligten Außenstandpunkt her beobachtbar bzw. messbar zu machen.

Die Entdeckung des
bedingten Reflexes

Etwa zur gleichen Zeit, gegen Ende des 19. Jahrhunderts, wurden von dem russischen Physiologen Iwan Pawlow (1849–1936) in St. Petersburg Untersuchungen über die Sekretion der Verdauungssäfte durchgeführt. Dabei bestätigte sich, was schon viele Hundeliebhaber vor Pawlow gesehen hatten: dass die Speicheldrüsen ihr Sekret nicht erst abgeben, wenn sich das Futter im Maul eines Hundes befindet, sondern bereits beim Anblick des Futters. Aber Pawlows Entdeckung ging darüber hinaus: Pawlow zeigte, dass auch ein neutraler Sinnesreiz, der mit dem Futter gar nichts zu tun hat, zum Auslöser der Speichelsekretion wird, nachdem er mehrmals gleichzeitig mit der Fütterung aufgetreten ist.

- Ertönt z.B. bei der Fütterung des Hundes jedesmal ein Glo-

ckenton, so beginnt der Speichel nach einiger Zeit auch zu fließen, wenn der Glockenton ohne Futter erklingt.

- Im Unterschied zu den angeborenen, durch den natürlichen Reiz ausgelösten Reflexen (hier: Speichelfluß als Reflex auf die Wahrnehmung des Futters) bezeichnete Pawlow solche Reflexe, die durch einen vormals neutralen Stimulus ausgelöst werden, als bedingte Reflexe.

Bedingte Reflexe sind objektiv feststellbar und unter Anwendung physikalischer Messmethoden quantitativ erfassbar. So kann die Stärke des bedingten Reflexes z.B. an der Zahl der Speicheltropfen oder an der in Sekunden gemessenen Latenz des Reflexes gemessen werden (vgl. Tab. 1.1.2). Als Latenz bezeichnet man dabei die Dauer der Zeitspanne zwischen Stimulus und Reaktion.

H	A	L
1	0	–
9	18	15
15	30	4
31	65	2
41	69	1
51	64	2

H = Häufigkeit, mit der Ton und Fütterung gleichzeitig erfolgt waren
 A = Anzahl der Speicheltropfen innerhalb der ersten 30 Sekunden nach dem Ton
 L = Latenz des Reflexes in Sekunden

Tab. 1.1.2: Stärke des bedingten Reflexes auf einen Ton von 637,5 Hertz (zit.n. Rohracher, 1963, 277)

Mit seinen Arbeiten schuf Pawlow die Grundlage für eine der bedeutendsten Forschungsrichtungen innerhalb der Lernpsychologie, die man heute als klassische Konditionierung bezeichnet.¹ Im Verlauf ihrer Entwicklung wurde der Begriff des bedingten Reflexes, der nur einzelne Organreaktionen (z.B. Speichelsekretion) umfaßt, zur bedingten Reaktion ausgeweitet, die auch komplexere Verhaltensweisen mit einschließt. In Anlehnung an den amerikanischen Sprachgebrauch spricht man heute allgemein von einer konditionierten Reaktion (CR). Entsprechend bezeichnet man den vormals neutralen Stimulus (NS) (z.B. den Glockenton), sobald er zum Auslöser der konditionierten Reaktion geworden ist, als konditionierten Stimulus (CS). Der natürliche Reiz, der die Reaktion als biologische Wirkung auslöst (z.B. das Futter), wird dagegen als unkonditionierter Stimulus (US) bezeichnet, die Reaktion darauf als unkonditionierte Reaktion (UR). Unter Anwendung dieser

Die klassische Konditionierung

1. Vgl. hierzu weiterführend: Spada, Ernst & Ketterer (1990, 327ff).

Terminologie läßt sich das Prinzip des klassischen Konditionierens wie folgt darstellen:

Die klassische Konditionierung

Wird ein unkonditionierter Stimulus (US), der eine spezifische, unkonditionierte Reaktion (UR) auslöst, regelmäßig mit einem neutralen Stimulus (NS) gekoppelt, so wird der ursprünglich neutrale Stimulus (wenn man die Kopplung nur oft genug durchgeführt hat) schließlich zu einem konditionierten Stimulus (CS), der diese Reaktion nun auch eigenständig (d.h. ohne Paarung mit dem unkonditionierten Stimulus) – als konditionierte Reaktion (CR) – auslöst.

Die operante Konditionierung

Gleichzeitig mit Pawlow berichtete Edward Lee Thorndike (1874–1949) über Experimente, mit denen er die Intelligenz von Tieren messen wollte. Dabei stellte er fest, dass z.B. Katzen, die in einen Käfig gesperrt waren, den man von innen öffnen konnte, den Öffnungsmechanismus schließlich durch Versuch und Irrtum fanden und von da an die erfolgreiche Verhaltensweise beibehielten. Von Thorndike später als Gesetz des Effektes formuliert, wurde diese Entdeckung zur Grundlage einer weiteren – nicht minder bedeutenden – Forschungsrichtung innerhalb der Lernpsychologie, die man heute als operante Konditionierung bezeichnet.¹

Während das Versuchstier bei der klassischen Konditionierung lernt, einen ursprünglich neutralen Stimulus mit einem Verhalten zu verbinden, das ursprünglich als natürliche Reaktion des Organismus auf einen anderen Stimulus eintritt, liegt hier eine andere Form des Lernens vor, bei dem neue Verhaltensweisen aufgrund ihrer positiven Konsequenzen erworben werden.

Seit Thorndikes ersten Versuchen wurden viele "Problem-Käfige" konstruiert, die dem Versuchstier die Möglichkeit bieten, durch Versuch und Irrtum eine erfolgreiche Verhaltensweise zu finden. Berühmt wurde die von B.F. Skinner entwickelte Skinner-Box, in der sowohl klassische als auch operante Konditionierungsexperimente durchgeführt werden können (vgl. Abb. 1.1.1).

1. Zur Entwicklung des Konzepts der operanten Konditionierung vgl. Coleman (1981).

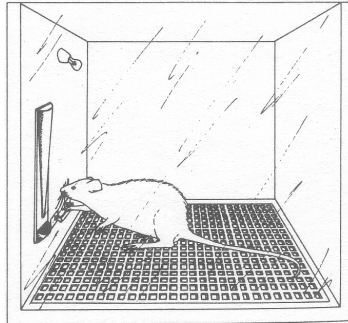


Abb. 1.1.1: Eine Skinner-Box für Versuche mit Ratten. An einer Wand ist ein Hebel angebracht. Drückt das Versuchstier den Hebel, so bewirkt dies eine kurzzeitige Öffnung des Futtermagazins. Darüber befindet sich eine Lampe, die Konditionierung auf Lichtreize ermöglicht. Der Boden besteht aus einem elektrifizierbaren Gitter, über das die Ratte Schmerzreizen ausgesetzt werden kann (nach Spada et al., 1990, 332)

Bei der folgenden Darstellung der operanten Konditionierung stützen wir uns auf Thorndikes (1913) erste Formulierung des Gesetzes:

Das Gesetz des Effektes

Die operante Konditionierung

- (1) Organismen reagieren auf neue Problemsituationen zunächst so lange probeweise mit instinktiven oder bereits früher erlernten Verhaltensweisen, bis eine der Reaktionen zufällig zu einem Erfolg führt.
- (2) Verknüpfungen von Situationen und Reaktionen, die von einem befriedigenden Gesamtzustand (satisfying state of affairs) begleitet sind, werden verstärkt. Die Stärke von Verknüpfungen, die von einem unbefriedigenden Gesamtzustand (annoying state of affairs) gefolgt sind, nimmt dagegen ab (Gesetz des Effektes).

Thorndikes Modellvorstellung des Lernens am Erfolg durch Versuch und Irrtum stand dabei in bewußtem Gegensatz zu der am Ende des 19. Jahrhunderts noch weit verbreiteten Meinung, dass Tiere in derartigen Situationen durch Denken zur Lösung kommen. In der weiteren Geschichte der Psychologie wurden Pawlows bedingter Reflex und Thorndikes Gesetz des Effektes zu den beiden Fundamenten, auf welche der von McDougall (1912) und Watson (1913) begründete Behaviorismus sich rasch zur vorherrschenden Richtung innerhalb der amerikanischen Psychologie entwickelte. Diese Position behielt er bis in die 70er Jahre des 20. Jahrhunderts bei.

Der Behaviorismus

In ideologischer Hinsicht war der Behaviorismus durch eine demokratische Grundauffassung von der Chancengleichheit aller Menschen geprägt, die davon ausgeht, dass die Verhaltensweisen von Mensch und Tier in erster Linie durch Umwelteinflüsse geprägt werden. Auf Grundlage der Prinzipien des klassischen und operanten Konditionierens konnte diese Auffassung dahingehend zugespitzt werden, dass es nur sehr wenige angeborene, aber sehr viele konditionierte Verhaltensweisen gibt. Alles, was ein Mensch denkt, fühlt und tut, ist nach Watson zum größten Teil gelernt, d.h. durch Wiederholung von Situationen und Reaktionen entstanden. In dieser Grundüberzeugung ging Watson (1931, zit.n. Rohracher, 1963, 280) so weit, dass er behauptete, "dass es keine Vererbung von Fähigkeiten, Talent, Temperament, seelischer Konstitution und Charaktereigenschaften gibt. Auch diese Dinge hängen von der Übung ab, die schon in der Wiege beginnt".

Orientierung am Vorbild der Physik

In methodologischer Hinsicht war der Behaviorismus durch den Versuch geprägt, ein strenges Kriterium für Wissenschaftlichkeit in der Psychologie durchzusetzen, wobei er sich am Vorbild der Physik orientierte, die allerdings nur in positivistischer Verkürzung rezipiert wurde. Das Mittel der Selbstbeobachtung wurde als unwissenschaftlich verboten (vgl. Janich, 1980); subjektseitig definierte Sachverhalte wie Gefühle, Gedanken oder Intentionen etc. wurden als "private events" (Buss, 1961) denunziert, die einer wissenschaftlichen Behandlung nicht zugänglich seien. Als Gegenstand der Psychologie kämen nur die objektseitig definierten und distanziert beobachtbaren Stimuli und Reaktionen in Betracht.

Das Scheitern des Behaviorismus

Dies war zwar konsequent gedacht, aber wegen der Besonderheit des Forschungsgegenstandes der Psychologie von vornherein zum Scheitern verurteilt. Denn anders als die Physik hat es die Psychologie ja nicht nur mit der unbelebten Materie zu tun, und anders als die Biologie auch nicht nur mit dem lebenden Organismus. Welche Konsequenzen sich daraus für Methodologie und Forschungsmethoden der Psychologie ergeben, wird weiter unten diskutiert. Hier sei nur so viel vorweggenommen, dass der Behaviorismus seit der um die Mitte der 70er Jahre des 20. Jahrhunderts einsetzenden kognitiven Wende, insbesondere seit den vielbeachteten Arbeiten von Kahnemann & Tversky (1979), als überholt gilt.

Kontrollfragen

- Wann begann sich die Psychologie als eigenständige empirische Wis-

senschaft zu konstituieren?

- Wie hatte die Vermögenspsychologie bis dahin die psychischen Phänomene erklärt?
- Warum handelte es sich dabei um Pseudo-Erklärungen?
- Was versteht man unter Verdoppelung der Realität?
- Warum stellt die Abwendung von der Vermögenspsychologie einen entscheidenden Fortschritt in der Psychologiegeschichte dar?
- Wie definiert Rohracher den Gegenstand der Psychologie?
- Welcher grundlegende Unterschied zwischen der Psychologie und den klassischen Naturwissenschaften ergibt sich aus dieser Definition?
- Welche (ersten) methodologischen Forderungen ergeben sich aus der Auffassung der Psychologie als Wissenschaft?
- Warum bedarf die wissenschaftliche Psychologie einer geklärten Wissenschaftssprache?
- Was bedeutet das Wort ‚Terminologie‘?
- Welches sind die Hauptgebiete der psychologischen Methodenlehre?
- Worin besteht der Unterschied zwischen objektseitig und subjektseitig definierten Sachverhalten?
- Warum ist eine Polarisierung von naturwissenschaftlicher vs. geisteswissenschaftlicher Psychologie wenig sinnvoll?
- Wo, wann und von wem wurde das erste Institut für experimentelle Psychologie gegründet?
- Welche Auffassungen des Philosophen David Hume haben der empirischen Psychologie den Weg bereitet?
- Welche Arten der Erfahrung wurden von David Hume unterschieden?
- Beschreiben Sie die Experimente, die zur Entdeckung der (absoluten) Unterschiedsschwelle geführt haben!
- Was versteht man unter der absoluten Unterschiedsschwelle?
- Was versteht man unter der relativen Unterschiedsschwelle?
- Wie lautet das Weber'sche Gesetz?
- Wovon ist die relative Unterschiedsschwelle abhängig?
- Worin besteht die wissenschaftshistorische Bedeutung der Entdeckung des Weber'schen Gesetzes?
- Erläutern Sie anhand von Pawlows klassischem Experiment, was man unter einem bedingten Reflex versteht!
- Woran kann die Stärke des bedingten Reflexes in Pawlows Experiment gemessen werden?
- Was versteht man unter der Latenz des bedingten Reflexes?
- Erläutern Sie die Termini ‚neutraler Stimulus‘, ‚konditionierter Stimulus‘, ‚unkonditionierter Stimulus‘, ‚unkonditionierte Reaktion‘ und ‚konditionierte Reaktion‘!
- Worin besteht das Prinzip der klassischen Konditionierung?
- Erläutern Sie den Unterschied zwischen klassischer Konditionierung und operanter Konditionierung!
- Wie lautet das Gesetz des Effektes?
- Wodurch war das Psychologieverständnis des Behaviorismus geprägt?
 - (a) in ideologischer,
 - (b) in methodologischer Hinsicht?
- Warum war der Behaviorismus zum Scheitern verurteilt?

1.2

Wissenschaftliche Psychologie und psychologisches Alltagswissen

Das Unbehagen
gegenüber der
Psychologie

Wenn man jemandem erzählt, dass man Psychologe ist oder Psychologie studiert, dann stößt man nicht selten auf eine Abwehrreaktion der Art: "Oh – da muss ich aber aufpassen". Oder man darf sich gleich einen Witz anhören:

"Treffen sich zwei Psychologen. Sagt der eine: ‚Hallo, wie geht es mir?‘"

Manchmal ist es auch folgender Witz, der schon ein wenig mehr an psychologischen Kenntnissen voraussetzt:

"In Skinners Laboratorium treffen sich zwei Ratten. Sagt die eine: ‚Den Skinner habe ich gut konditioniert. Jedes Mal, wenn ich diesen Hebel drücke, gibt er mir eine Futterpille.‘"

Solche Abwehrreaktionen und Witze sind Ausdruck eines gewissen Unbehagens gegenüber Psychologen, die vermeintlich oder tatsächlich den Anspruch erheben, über unsere ureigensten Angelegenheiten besser Bescheid zu wissen als wir selbst. In den oben erzählten Witzen wird dieser Anspruch dann ins Lächerliche gezogen.

Interpretations-
monopol der
Psychoanalyse?

An der Entstehung dieses Vorbehaltes gegenüber der Psychologie und ihren Vertretern ist die Psychologie selbst nicht ganz unbeteiligt. Das gilt insbesondere für die Psychoanalyse, die mit dem Konzept des Unbewussten ja den Anspruch erhebt, gleichsam in die Tiefe unserer Seele hineinzublicken und dort Motive zu entdecken, derer wir uns selbst nicht bewusst sind – z.B. einen Ödipuskomplex: Todeswünsche gegenüber dem Vater und sexuelle Wünsche gegenüber der Mutter. Und wenn wir solche Wünsche von uns weisen, dann wird das als Widerstand ausgelegt, den wir der Bewusstmachung unserer unbewussten Motive entgegensetzen.¹

1. Zur Verwendungsweise der o.g. psychoanalytischen Termini vgl. Laplanche & Pontalis (1973).

Solchen vulgärfreudianischen Tendenzen in der Psychoanalyse, die dem Psychoanalytiker qua Autorität ein Interpretationsmonopol zusprechen und die Interpretationen des Analytikers gegen jegliche Kritik immunisieren (Fromm, 1979), stehen zwar einerseits im Gegensatz zu Freud selbst, der in seinen *Vorlesungen zur Einführung in die Psychoanalyse* (Freud, 1917) deutlich darauf hinweist, dass anhand von Indizien aufgestellte Behauptungen über unbewusste Motive auch tatsächlich – und unter Einbeziehung der Klienten – zu begründen sind. Andererseits hat Freud selbst diesen pseudowissenschaftlichen Tendenzen jedoch insofern das Feld bereitet, als er methodologische und methodische Ausführungen darüber, wie diese Begründungspflichten einzulösen wären, weitgehend schuldig blieb (vgl. Aschenbach, 1985).

Die genannten Vorbehalte gegenüber Psychologen sind aber nicht nur der Psychoanalyse anzulasten. Sie werden auch von manchen Lehrbüchern der Psychologie – sogar der psychologischen Methodenlehre – genährt, die eine grundsätzliche Überlegenheit wissenschaftlicher Theorien gegenüber dem Alltagswissen postulieren (z.B. Huber, 1987, 13ff) und/oder die Wissenschaftlichkeit einer Hypothese an der bloß schematischen Befolgung bestimmter Regeln festmachen (z.B. Bortz & Döring, 1995, 7f), die nicht weiter begründet werden als damit, dass sie eben wissenschaftlich seien. Jedoch gerade das, was Wissenschaft eigentlich ausmacht, geht darüber verloren:

"Wissenschaftlichkeit, das heißt zu wissen, was man weiß; unwissenschaftlich ist das dogmatische Wissen. Wissenschaftlich sein, das heißt, mit Gründen zu wissen; unwissenschaftlich ist das Hinnehmen fertiger Meinungen. Wissenschaftlich ist das Wissen mit dem Bewusstsein von den jeweils bestimmten Grenzen des Wissens; unwissenschaftlich ist alles Totalwissen, als ob man im Ganzen Bescheid wüsste. Wissenschaftlich ist grenzenlose Kritik und Selbstkritik; unwissenschaftlich ist die Besorgnis, der Zweifel könne lähmen." (Jaspers, 1945, zit.n. Kempf, 1986, 51)

Über die Methoden der Wissenschaft selbst ist damit noch nichts ausgesagt; wohl jedoch über die Kriterien, denen sie genügen müssen, damit sie als wissenschaftlich gelten können: Wissenschaftliche Methoden sind Instrumente des systematischen Zweifels an der Wahrheit von Behauptungen. Sie zwingen uns, für das, was wir zu wissen meinen, Gründe vorzubringen. Und sie nehmen uns zugleich in die Pflicht, die Grenzen unseres Wissens zu explizieren, d.h. insbesondere auch: seine impliziten Voraussetzungen transparent zu machen.

Grundsätzliche Überlegenheit der Wissenschaft?

Die Begründungsbedürftigkeit von Wissenschaft

Das praktische Fundament der Wissenschaft

Und dafür gibt es einen guten Grund, den die Wissenschaft mit dem Alltagsleben teilt. Es ist ja *nicht* so, wie Huber (1987, 17) meint, dass es bei der Anwendung alltagspraktischen Wissens "oft schlicht irrelevant" wäre, "ob das Wissen richtig oder falsch oder auch widersprüchlich ist". Im Gegenteil erfüllt das Alltagswissen eine wichtige Orientierungsfunktion für unser tägliches Handeln. Und wenn wir in unserem Handeln erfolgreich sein wollen, sind wir darauf angewiesen, dass es auch richtig ist. D.h. wir haben ein praktisches Interesse daran, wahre Aussagen von falschen zu trennen. Auch die Faszination, welche das Vorbild der Naturwissenschaften auf die Begründer der erfahrungswissenschaftlichen Psychologie ausgeübt hat – und bis heute ausübt –, rührt ja nicht einfach her aus ihrer Wissenschaftlichkeit per se, sondern aus der technischen Verwertbarkeit der neuzeitlichen Naturwissenschaften (vgl. Hoyningen-Huene, 1989b), die zu einer unglaublichen Erweiterung der Handlungsmöglichkeiten des Menschen geführt hat.

Auch die psychologische Wissenschaft kann in diesem Sinne als eine institutionalisierte Form der Wissensbildung verstanden werden, die einem praktischen Interesse dient, indem sie eine bestimmte Klasse von Sachverhalten *erklärt*. Diese Klasse von Sachverhalten ist es, die den Gegenstand der Psychologie ausmacht und die Psychologie von anderen Wissenschaften, wie z.B. der Biologie, unterscheidet.

Darüber, welche logische Struktur die Erklärungen haben sollen, ist damit noch nichts ausgesagt. Wir können aber schon so viel festhalten, dass (wissenschaftliche) Erklärungen in Form von Theorien dargestellt werden, wobei unter einer Theorie zunächst nichts anderes als ein mehr oder minder komplexes System von Aussagen zu verstehen ist. Als wissenschaftliche Aussagen kommen nur *wahrheitsfähige* Aussagen in Betracht. D.h. es muss festgelegt sein, wie für oder gegen die Geltung einer Aussage zu argumentieren ist bzw. wodurch sie bestätigt oder widerlegt werden kann.

Von den Aufgaben der Wissensbildung bis hin zu den Methoden der Geltungsprüfung von Theorien besteht ein systematischer Begründungszusammenhang. Denn: Welche Kriterien zur Beurteilung der Gültigkeit einer Theorie angemessen sind, hängt außer von der Bedeutung der Termini, die darin verwendet werden, auch von der logischen Struktur der Theorie ab; und welche logi-

sche Struktur eine Theorie haben muss, damit sie eine bestimmte Erklärungsaufgabe zu erfüllen vermag, hängt von eben dieser Aufgabe ab.

Kontrollfragen

- Welches praktische Interesse liegt dem Streben der Wissenschaften nach Wahrheit zugrunde?
- Was versteht man unter einer Theorie, und welchen Kriterien müssen die darin enthaltenen Aussagen genügen?
- Welcher Begründungszusammenhang besteht zwischen den Aufgaben der Wissensbildung, der logischen Struktur von Theorien und den Methoden ihrer Geltungsprüfung?

1.3

Propädeutische Grundlagen

Vorüberlegungen zu Terminologiebildung und Logik

Um diesen Begründungszusammenhang konkret herausarbeiten und einen systematischen Aufbau der psychologischen Methodenlehre leisten zu können, müssen wir jedoch zuerst einige Vorüberlegungen zu Fragen der Terminologiebildung und der Logik anstellen. Dabei beschränken wir uns zunächst auf die Klärung einiger Grundbegriffe, ohne die darin anklingenden erkenntnistheoretischen Fragen im einzelnen zu behandeln.¹

Diesen propädeutischen Vorüberlegungen kommt für die psychologische Methodenlehre in mehrfacher Hinsicht eine Schlüsselrolle zu: Wir benötigen sie,

- um den das ganze Buch durchziehenden Aufbau der psychologischen Terminologie leisten zu können,
- um den Unterschied zwischen der Psychologie und anderen Erfahrungswissenschaften herausarbeiten zu können,
- um die methodischen Schritte begründen zu können, die zur Geltungsprüfung psychologischer Theorien erforderlich sind,
- um die verschiedenen (naturwissenschaftlichen und kulturwissenschaftlichen) Paradigmen, welche innerhalb der Psychologie bestehen, miteinander integrieren und systematisch aufeinander beziehen zu können –
- und wir benötigen sie ganz allgemein, um Sprachverwirrung und daraus resultierende Fehlschlüsse zu vermeiden, wie sie z.B. dem folgenden Theorem zugrunde liegen, dessen Beweis freilich nur in englischer Sprache gelingt, weil die darin verwendeten Worte im Deutschen nicht jene Doppelbedeutung haben, auf deren Vermischung er beruht.

1. Zu einem systematischen Aufbau der logischen Propädeutik vgl. das Standardwerk von Kamlah & Lorenzen (1967) sowie weiterführend auch Janich, Kambartel & Mittelstraß (1974), Lorenzen & Schwemmer (1975) und Lorenzen (1987).

Theorem: The more you know, the less you earn

Sprachverwirrung
und Fehlschlüsse

Proof:

As all of us know

$$(1) \quad \text{Knowledge} = \text{Power}$$

and

$$(2) \quad \text{Time} = \text{Money}.$$

As any engineer knows

$$(3) \quad \text{Power} = \text{Work/Time}.$$

Substituting (3) into (1) we get

$$(4) \quad \text{Knowledge} = \text{Work/Time}$$

and substituting (2) into (4) we get

$$(5) \quad \text{Knowledge} = \text{Work/Money}.$$

Solving for Money results in

$$(6) \quad \text{Money} = \text{Work/Knowledge}$$

and accordingly

$$(7) \quad \lim_{\text{Knowledge} \rightarrow \infty} (\text{Money}) = 0.$$

Quod erat demonstrandum.

1.3.1 Beobachtungssprache und Theoriesprache

Beim Aufbau einer wissenschaftlichen Terminologie kann zwischen verschiedenen Sprachebenen unterschieden werden, deren erste die Beschreibungssprache ist. Beschreibungssprachliche Termini sind solche, deren Gebrauch man anhand von Beispielen und Gegenbeispielen einüben kann, z.B. die Wörter ‚Futter‘, ‚Speichelfluss‘ und ‚Glockenton‘, die wir zur Beschreibung von Pawlows Experimenten verwendet haben. Viele dieser Wörter können wir relativ problemlos aus der natürlichen Sprache übernehmen. Andere sind strittig und bedürfen zwecks Eindeutigkeit des wissenschaftlichen Sprachgebrauchs einer expliziten Vereinbarung.

Beschreibungs-
sprache

Aussagen, Fragen und Aufforderungen

Indem wir die Einübung beschreibungssprachlicher Termini rekonstruieren, gewinnen wir zugleich erste syntaktische Unterscheidungen, d.h. erste Termini, in denen wir über die Sprache und die Regeln ihres Aufbaus sprechen können, z.B. die Unterscheidung zwischen verschiedenen Arten von Sätzen: Aussagen (A), Fragen (A?) und Aufforderungen (A!). Diese Unterscheidung wird nicht durch eine deduktive Definition im Sinne einer Wortsetzungsregel eingeführt, sondern vielmehr redend und handelnd eingeübt, d.h. indem Reden mit nichtsprachlichen Handlungen verbunden werden. Ob es sich bei einem Satz um eine Aussage ("Das Kind ist ruhig."), um eine Frage ("Ist das Kind ruhig?") oder um eine Aufforderung ("Beruhige das Kind!") handelt, hängt davon ab, welcher *Art* die angemessene Antwort darauf ist; und ob die darin vorkommenden Wörter richtig verstanden wurden, zeigt sich an der Angemessenheit ihres *Inhalts*. Ob z.B. die Wortfolge "Du gehst" als Aussage, Frage oder Aufforderung zu klassifizieren ist, ergibt sich aus dem Handlungszusammenhang zwischen Sprecher und Angesprochenem: So kann z.B. der Angesprochene einer Aussage zustimmen oder sie bestreiten; eine Frage bejahen oder verneinen und eine Aufforderung befolgen oder sich weigern. Befolgt er eine Aufforderung, so zeigt sich das richtige Verständnis der darin vorkommenden Wörter wiederum daran, dass er auch tatsächlich das tut, wozu ihn der Sprecher aufgefordert hat usw. Im Anschluss an den Sprachpsychologen Karl Bühler (1934) kann man daher auch sagen, dass der Gebrauch dieser ersten Wörter einer Wissenschaftssprache *empirisch* kontrolliert wird.

Elementaraussagen

Die einfachste Aussage ist die Elementaraussage, z.B. "Das Kind ist ruhig". Sie setzt sich aus zwei Bestandteilen zusammen: Dem Nominativ 'das Kind' (N), der einen bestimmten Gegenstand bezeichnet, und dem grammatikalischen Prädikat "ist ruhig". Letzteres besteht aus dem Prädikator 'ruhig' (P), der dem Gegenstand zugesprochen wird, und dem Bindewort "ist", der Kopula (ϵ). Formal:

$$N \ \epsilon \ P$$

Eine solche Elementaraussage, bei der einem Gegenstand ein Prädikator zugesprochen wird, bezeichnet man als affirmative Aussage. Im Unterschied dazu wird bei einer negativen Aussage dem Gegenstand ein Prädikator abgesprochen, z.B. "Das Kind ist nicht ruhig". Formal:

N ε' P

Als Elementaraussagen bezeichnet man Aussagen dieser Art deshalb, weil sie noch nicht logisch zusammengesetzt sind.

Der wesentliche Unterschied zwischen Nominatoren und Prädikatoren besteht darin, dass Nominatoren genau einen Gegenstand bezeichnen. Prädikatoren können dagegen vielen Gegenständen zugesprochen werden.

Nominatoren und Prädikatoren

Bei den Nominatoren kann man unterscheiden zwischen Eigennamen, z.B. ‚Hans‘, und Kennzeichnungen, z.B. ‚das Kind‘. Wichtig ist dabei die Eindeutigkeit, mit der der gemeinte Gegenstand identifiziert wird.

Berichten wir z.B. über eine entwicklungspsychologische Studie an mehreren Kindern, so ist die Kennzeichnung ‚das Kind‘ nicht geeignet, um ein bestimmtes Kind zu identifizieren. Wir werden dann eine komplexere Kennzeichnung vornehmen müssen, z.B. ‚jenes Kind, das im Eignungstest versagt hatte‘. Eine solche Kennzeichnung ist aber ihrerseits schon logisch zusammengesetzt und daher nicht elementar.

Auch bei den Prädikatoren lassen sich verschiedene Arten unterscheiden. So sind z.B. in der folgenden, relativ einfachen psychologischen Aussage drei verschiedene Arten von Prädikatoren enthalten: "Das Kind ist *unruhig*, weil es sich *unbeachtet wähnt* und *Aufmerksamkeit sucht*".

Beobachtungs-, Deutungs- und Handlungsprädikatoren

Diese verschiedenen Arten von Prädikatoren, mit denen wir es in der psychologischen Wissenschaftssprache zu tun haben, erfordern sowohl unterschiedliche Schritte zur Klärung ihrer Bedeutung als auch unterschiedliche Schritte zur Überprüfung der Geltung von Aussagen, die mit ihnen gebildet wurden.

Ob ein Kind ruhig ist oder nicht, kann unmittelbar beobachtet werden; nicht mehr jedoch, dass es Aufmerksamkeit sucht. Und dass es Aufmerksamkeit sucht, *weil* es sich unbeachtet wähnt, geht erst recht über das hinaus, was der direkten Beobachtung zugänglich ist.

Während ‚unruhig sein‘ ein Beobachtungspredikator ist, also der Beobachtungssprache angehört, stellt ‚sich unbeachtet wähnen‘ einen Deutungsprädikator dar, der der Theoriesprache angehört. Es handelt sich dabei, wie man in der Psychologie auch zu sagen pflegt, um ein theoretisches Konstrukt, das zum beobachtbaren

Verhalten nach bestimmten Regeln hinzukonstruiert wird und einen hypothetischen Charakter hat: dass sich das Kind ‚unbeachtet wähnt‘, ist eine Annahme, die vom Außenstandpunkt eines neutralen Beobachters nicht definitiv bewiesen werden kann. ‚Aufmerksamkeit suchen‘ wiederum ist ein Handlungsprädikator, der bereits eine theoriehaltige Beschreibung des beobachteten Verhaltens leistet.

Exemplarische Einführung von Beschreibungsprädikatoren

Die Beobachtungsprädikatoren sind die einfachsten Termini einer Wissenschaftssprache. Ihre Bedeutung kann man letztlich nur so klären, dass man ihren Gebrauch an Beispielen und Gegenbeispielen einübt. Man spricht dann von einer exemplarischen Einführung der Prädikatoren: Indem einer Menge von Beispielen der Prädikator zugesprochen und einer Menge von Gegenbeispielen der Prädikator abgesprochen wird, kann man z.B. einüben, was der Unterschied ist zwischen ‚ruhig sein‘ und ‚nicht ruhig sein‘, zwischen ‚schreien‘ und ‚nicht schreien‘, zwischen ‚zittern‘ und ‚nicht zittern‘ usw.

Theoriehaltigkeit von Handlungsbeschreibungen

Den Gebrauch von Handlungsprädikatoren kann man zwar auch exemplarisch einüben (sie gehören daher ebenfalls der Beschreibungssprache an), z.B. ‚spielen‘, ‚arbeiten‘, ‚lernen‘, bei der Beschreibung von Handlungen tritt aber zur Beschreibung eines beobachtbaren Verhaltens noch eine Deutung dieses Verhaltens hinzu. Ob ein Kind ‚spielt‘ oder ‚lernt‘, lässt sich aufgrund von Beobachtung allein noch nicht entscheiden. Als ‚Lernen‘ werden wir sein Tun vielmehr erst dann identifizieren, wenn es zum Zwecke des Wissenserwerbes geschieht etc.

Theoretische Konstrukte

Zur Theoriesprache gehören auch die Termini, die wir zur Darstellung der klassischen Konditionierung verwendet haben (unkonditionierter Stimulus, unkonditionierte Reaktion, neutraler Stimulus). Im Unterschied zu den beobachtungssprachlichen Prädikatoren (‚Futter‘, ‚Speichelfluss‘, ‚Glockenton‘), mit denen wir Pawlows Experimente beschrieben hatten, setzt die Verwendung dieser Termini bereits das Wissen um eine empirische Gesetzmäßigkeit voraus:

- Dass das Futter ein US (bezüglich des Speichelflusses) ist, bedeutet, dass dieser Stimulus regelmäßig die Reaktion ‚Speichelfluss‘ auslöst, und
- dass der Speichelfluss eine UR (auf das Futter) ist, bedeutet umgekehrt, dass er durch den Stimulus ‚Futter‘ regelmäßig ausgelöst wird.

- Dass der Glockenton ein NS (bezüglich des Speichelflusses) ist, bedeutet dagegen, dass dieser Stimulus, die Reaktion ‚Speichelfluss‘ nicht auslöst.

Dass es sich bei diesen Termini um theoretische Konstrukte handelt, schlägt sich auch im typischen Aufbau klassischer Konditionierungsexperimente nieder, der in drei Phasen erfolgt:

Das klassische Konditionierungsexperiment (nach Spada et al., 1990, 327f)

Das klassische
Konditionierungs-
experiment

Kontrollphase: Zunächst ist zu prüfen, ob der US ein zuverlässiger Auslöser der UR ist. Der NS darf die UR in der Kontrollphase dagegen noch nicht auslösen. Er ruft jedoch – falls er dem Tier unbekannt ist – eine sog. Orientierungsreaktion (OR; Sokolov, 1963) hervor. Sie beinhaltet eine Hinwendung zur Reizquelle und eine allgemeine Steigerung der Aufmerksamkeit, die sich beispielsweise in einer Pupillenerweiterung und einer Erhöhung des Herzschlages ausdrücken kann. Die UR sollte sich von der OR deutlich unterscheiden, da andernfalls die OR fälschlich als CR interpretiert werden könnte (Gormezano & Moore, 1969).

Konditionierungsphase: Der NS und der US werden mehrmals gemeinsam dargeboten, wobei der US die UR auslöst. Sehr wichtig ist die zeitliche Beziehung zwischen den beiden Reizen (Jones, 1962; Mackintosh, 1974). Untersuchungen haben gezeigt, dass die verzögerte Konditionierung (bei der der US nach dem NS einsetzt, der NS aber während des US noch andauert) am raschesten zum Aufbau einer konditionierten Reaktion führt. Simultane Konditionierung (bei der US und NS gleichzeitig dargeboten werden) und Spurenkonditionierung (bei der der NS dem US vorausgeht) liefern etwas schlechtere Ergebnisse. Bei rückwirkender Konditionierung (bei der der US dem NS vorausgeht) ist umstritten, ob überhaupt ein systematischer Lerneffekt eintritt (Jones, 1962; Spetch, Wilkie & Pinel, 1981).

Löschungsphase: Der CS wird ohne Kopplung mit dem US dargeboten. Im Falle einer erfolgreichen Konditionierung tritt die CR in den ersten Durchgängen deutlich auf, im Allgemeinen jedoch, ohne die Stärke der UR zu erreichen. Nach einer Reihe von Darbietungen nimmt die Stärke der CR immer weiter ab, bis sie gelöscht ist.

Operationalisierung	Die Kontrollphase des Experimentes dient dabei dazu, die theorie-sprachlichen Termini (z.B. NS) in beobachtungssprachliche Termini (z.B. Glockenton) zu übersetzen. Diese Übersetzung theoretischer Konstrukte in beobachtbare Sachverhalte bezeichnet man allgemein als Operationalisierung. Man sagt dann auch z.B.: Der NS wird durch einen Glockenton operationalisiert.
Mehrstellige Prädikatoren	Der Beobachtungsprädikator ‚Glockenton‘ unterscheidet sich von dem theoretischen Konstrukt ‚neutraler Stimulus‘ noch in einer anderen Hinsicht: Während es sich beim ‚Glockenton‘ um einen einstelligen Prädikator handelt, der genau einem Gegenstand zu- oder abgesprochen wird, handelt es sich beim neutralen Stimulus um einen zweistelligen Prädikator. Neutral ist der NS nur in Bezug auf eine ganz bestimmte Reaktion (die UR, die durch den Konditionierungsvorgang in eine CR verwandelt werden soll). D.h. der Terminus ‚neutral‘ stellt eine bestimmte Beziehung zwischen einem bestimmten Stimulus und einer bestimmten Reaktion her, oder, wie wir auch sagen können, er setzt den Stimulus und die Reaktion in eine bestimmte Relation zueinander. Solche Prädikatoren, die zwei oder mehrere Gegenstände zueinander in Beziehung setzen, bezeichnet man als mehrstellige Prädikatoren oder als Relatoren, z.B. die zweistelligen Prädikatoren ‚jemanden anschreien‘ (Beobachtungsprädikator), ‚miteinander streiten‘ (Handlungsprädikator), ‚mit jemandem einen Konflikt haben‘ (Deutungsprädikator); oder die mehrstellige Relation ‚zwischen‘, die stets drei (oder mehr) Gegenständen zu- oder abgesprochen wird, z.B. "der Mediator vermittelt zwischen den (mindestens zwei) Konfliktparteien".
Ordnungsrelatoren	<p>Eine besondere Form der mehrstelligen Prädikatoren sind die (zweistelligen) Ordnungsrelatoren, die die Grundlage für die Einführung quantitativer Begriffe und für die Messung von Quantitäten bilden.</p> <p>Ein zweistelliger Prädikator (z.B. ‚länger als‘) definiert genau dann eine Ordnungsrelation, wenn er die folgenden Eigenschaften aufweist:</p> <ul style="list-style-type: none"> • <i>Irreflexivität</i>: N ist nicht länger als N • <i>Asymmetrie</i>: Wenn N_1 länger als N_2 ist, dann ist N_2 nicht länger als N_1 • <i>Transitivität</i>: Wenn N_1 länger ist als N_2 und N_2 länger ist als N_3, dann ist N_1 länger als N_3

Erst wenn der Relator ‚länger als‘ diese Eigenschaft erfüllt, kann man eine Ordinalskala (Rangskala) konstruieren, die jedem Gegenstand N_i eine Maßzahl x_i zuordnet, die die Länge der Gegenstände misst, indem sie die ‚länger als‘-Relation zwischen den Gegenständen auf die ‚größer als‘-Relation zwischen den Maßzahlen abbildet.

Erst dadurch, dass man (zumindest) eine Ordinalskala der Länge angeben kann, ist der Begriff der Länge als quantitative Dimension definiert. D.h. der Terminus ‚Länge‘ bezeichnet die Dimension, auf der die Messobjekte mittels der ‚länger als‘-Relation angeordnet werden.

Definition
quantitativer
Dimensionen

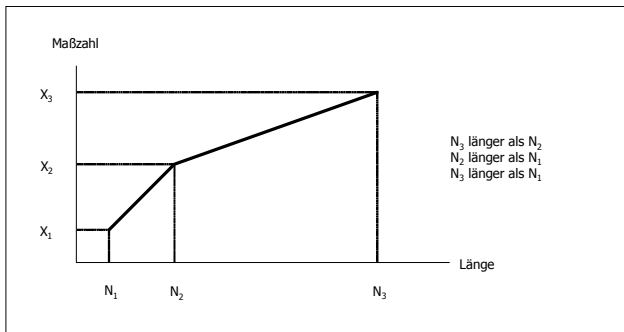


Abb. 1.3.1: Ordinale Längenmessung

In der Psychologie nimmt man es damit oft nicht so genau und verwendet quantitative Begriffe wie ‚Intelligenz‘, ohne dafür wirklich eine Skala anzugeben. Stattdessen werden diese Begriffe lediglich durch bestimmte Quantifizierungsverfahren (z.B. Intelligenztests) operationalisiert. Man spricht dann von einer Messung per fiat. Auf dieses Problem wird im Zusammenhang mit der Diskussion der Messung in der Psychologie noch genauer einzugehen sein.

Messung per fiat

1.3.2 Terminologische Regeln

Hat man erste Prädikatoren einer Wissenschaftssprache bereits exemplarisch eingeübt, so kann man ihre Verwendungsweise durch Prädikatorenregeln zusätzlich absichern. Auch die natürliche Sprache folgt solchen Regeln, wie z.B. der, dass Junggesellen unverheiratet sind. Der wesentliche Unterschied zur Alltagssprache besteht aber nun darin, dass diese Regeln nicht nur sprach-

Prädikatorenregeln

gebräuchlich befolgt, sondern explizit vereinbart werden. Z.B. können wir festhalten: "Eigennamen sind Nominatoren", bzw. formal: $[N \varepsilon \text{ Eigename}] \Rightarrow [N \varepsilon \text{ Nominator}]$. Das Zeichen " \Rightarrow " ist dabei der Regelpfeil und darf immer nur in Pfeilrichtung gelesen werden: $[N \varepsilon \text{ Eigename}]$ "erlaubt" $[N \varepsilon \text{ Nominator}]$. Allgemein hat eine Prädikatorenregel die Form

$$[N \varepsilon P] \Rightarrow [N \varepsilon Q].$$

Sie besagt: "Wer $[N \varepsilon P]$ behauptet, darf $[N \varepsilon Q]$ nicht bestreiten". Wenn wir z.B. bei der Analyse der Elementaraussage "Hans ist ruhig" das Wort ‚Hans‘ als Eigennamen identifiziert haben, dürfen wir nicht gleichzeitig bestreiten, dass das Wort ‚Hans‘ in dieser Aussage die Stelle des Nominators einnimmt.

Verteidigung von
Elementaraussagen

Solange Beschreibungsprädikatoren bloß exemplarisch bestimmt sind, kann die Verteidigung einer Aussage der Form $[N \varepsilon P]$ nur durch den Vergleich mit der Liste von Beispielen und Gegenbeispielen erfolgen. Findet sich der Gegenstand noch in keiner der beiden Listen, so ist neu zu entscheiden, ob eine der Listen um den fraglichen Gegenstand zu erweitern ist oder ob zu seiner Beschreibung die Einführung eines neuen Prädikators angemessen wäre.

Mit der expliziten Vereinbarung von Prädikatorenregeln eröffnet sich uns dagegen erstmals die Möglichkeit, für die Wahrheit oder Falschheit einer Aussage zu argumentieren, z.B.:

$$\frac{["Hans" \varepsilon \text{ Eigename}] \Rightarrow [N \varepsilon \text{ Nominator}]}{["Hans" \varepsilon \text{ Nominator}]}$$

Gilt die Regel

$$[N \varepsilon P] \Rightarrow [N \varepsilon Q],$$

dann gilt stets auch die Regel

$$[N \varepsilon' Q] \Rightarrow [N \varepsilon' P].$$

Ist z.B. sichergestellt, dass das Wort ‚ruhig‘ in der Aussage "Hans ist ruhig" kein Nominator ist, dann ist aufgrund obiger Regel auch sichergestellt, dass ‚ruhig‘ kein Eigename ist bzw. genauer: dass das Wort ‚ruhig‘ in dieser Aussage nicht als Eigename verwendet wird.

Gelten gleichzeitig die Regeln

$$[N \varepsilon P] \Rightarrow [N \varepsilon Q]$$

und

$$[N \varepsilon Q] \Rightarrow [N \varepsilon P],$$

so sind die Prädikatoren P und Q inhaltsgleich, d.h. sie stellen denselben Sachverhalt dar. Wir können dann auch schreiben:

$$[N \varepsilon P] \Leftrightarrow [N \varepsilon Q].$$

Solche inhaltsgleichen Prädikatoren unterscheiden sich nur hinsichtlich ihrer Lautgestalt, wie z.B. die oben verwendeten Termini ‚Stimulus‘ und ‚Sinnesreiz‘, die dasselbe bedeuten.¹

Bereits oben haben wir von einer Prädikatorenregel Gebrauch gemacht, wonach Eigennamen keine Prädikatoren sind. Diese Regel hatte die Form

$$[N \varepsilon P] \Rightarrow [N \varepsilon' Q]$$

bzw. in Worten $[N \varepsilon P]$ "verbietet" $[N \varepsilon Q]$, und besagt: "Wer $[N \varepsilon P]$ behauptet, darf nicht gleichzeitig auch $[N \varepsilon Q]$ behaupten". Gilt diese Regel, so sagt man auch: P und Q sind miteinander unverträglich bzw. sie stehen in einem konträren Gegensatz zueinander.

Gilt die Regel

$$[N \varepsilon P] \Rightarrow [N \varepsilon' Q],$$

dann gilt stets auch die Regel

$$[N \varepsilon Q] \Rightarrow [N \varepsilon' P].$$

Indem wir z.B. festgelegt haben, dass Eigennamen keine Prädikatoren sind, haben wir auch bereits festgelegt, dass Prädikatoren keine Eigennamen sind.²

Gelten gleichzeitig die Regeln

$$[N \varepsilon P] \Rightarrow [N \varepsilon' Q]$$

Inhaltsgleichheit von Prädikatoren

Unverträglichkeit von Prädikatoren

Kontradiktorischer Gegensatz

-
1. Die Aussagen "Die Termini P und Q sind inhaltsgleich.", "Die Termini P und Q stellen denselben Sachverhalt dar." und "Die Termini P und Q haben dieselbe Bedeutung." sind also ebenfalls wieder inhaltsgleiche Aussagen.
 2. Damit ist freilich nicht ausgeschlossen, dass ein Wort mit derselben Lautgestalt, z.B. "Müller", einmal als Eigenname und ein anderes Mal als Prädikator verwendet werden kann, wie z.B. in dem Satz "Hans ist Müller".

und

$$[N \varepsilon' Q] \Rightarrow [N \varepsilon P],$$

so stehen die beiden Prädikatoren in kontradiktorischem Gegensatz zueinander. Man sagt dann auch: "P ist das Gegenteil von Q", z.B. ‚farbig‘ und ‚farblos‘. Wir können die beiden Regeln zusammenfassen zu:

$$[N \varepsilon P] \Leftrightarrow [N \varepsilon' Q].$$

Polar konträrer
Gegensatz

Ein weiterer Spezialfall des konträren Gegensatzes ist der polar-konträre Gegensatz, wie er z.B. zwischen den Prädikatorenpaaren ‚klein – groß‘, ‚hell – dunkel‘ oder auch ‚intelligent – unintelligent‘ besteht, die gleichsam die Endpole einer Skala bezeichnen, auf der es einen kontinuierlichen Übergang von einem Pol zum anderen gibt. Polar-konträre Prädikatoren, z.B. ‚kurz‘ und ‚lang‘, erfordern zu ihrer Einführung daher vier Schritte:

- erstens die Einführung eines zweistelligen Ordnungsrelators ‚länger als‘ und
- (zweitens) anhand derselben Beispiele und Gegenbeispiele die Einführung seines Gegenteils ‚kürzer als‘, so dass
 $[N1 \text{ länger als } N2] \Leftrightarrow [N2 \text{ kürzer als } N1];$
- drittens die Konstruktion einer (ordinalen) Längenskala und schließlich
- (viertens) die Bezeichnung der Skalenpole mit ‚kurz‘ und ‚lang‘.

Definitionen

Wie wir bereits gesehen haben, dienen Prädikatorenregeln der Präzisierung der Verwendungsweise von bereits exemplarisch eingeübten Prädikatoren. Im Unterschied dazu dienen Definitionen zur Erhöhung der Sprachökonomie, und zwar durch die Einführung neuer Termini, die es erlauben, komplexe Sachverhalte mit einem einzigen Prädikator auszudrücken. Statt z.B. davon sprechen zu müssen, dass bei einem gegebenen Konditionierungsexperiment "der US nach dem NS einsetzt (P) und der NS während des US noch andauert (Q)", können wir aufgrund der oben gegebenen Definition auch kurz und bündig sagen, dass es sich in dem Experiment um eine verzögerte Konditionierung (R) handelt.

Formal gesehen, handelt es sich bei der Definition um eine terminologische Regel, die in beiden Richtungen zu lesen ist, z.B.

$$[N \varepsilon R] \Leftrightarrow [N \varepsilon P] \text{ und } [N \varepsilon Q].$$

Auf der linken Seite des Definitionspfeiles steht dabei eine Elementaraussage (z.B. $[N \in R]$), die man auch als Definiendum (das zu Definierende) bezeichnet. Im Unterschied dazu steht auf der rechten Seite des Definitionspfeiles eine logisch zusammengesetzte Aussage von beliebiger Komplexität (in unserem Beispiel: $[N \in P]$ und $[N \in Q]$), die als Definiens (das Definierende) bezeichnet wird. Dass man in beiden Richtungen von einer Aussage zur anderen übergehen darf, bedeutet schließlich, dass Definiens und Definiendum inhaltsgleich sind und somit denselben Sachverhalt darstellen.

Indem wir die Termini einer Wissenschaftssprache ausdrücklich vereinbaren, werden einerseits terminologische Regeln explizit gemacht, nach denen wir uns sprachgebräuchlich ohnedies bereits gerichtet hatten (z.B. "Junggesellen sind unverheiratet"). Zum anderen kann es aber auch geschehen, dass implizite Regeln der natürlichen Sprache abgeändert werden. So wird z.B. in der Zoologie der Wal zu den Säugetieren gerechnet, während die Umgangssprache von Walfischen spricht.

Erfahrungsbedingtheit terminologischer Regeln

Obwohl man dabei zwar einerseits darauf achten muss, dass die logische Korrektheit (d.h. Widerspruchsfreiheit) der Prädikatenregeln und Definitionen sichergestellt ist, gibt es andererseits keinen logisch zwingenden Grund, warum die Bedeutung eines bestimmten Terminus so und nicht anders zu vereinbaren ist. Terminologische Regeln sind auch keine empirischen Hypothesen, die man als richtig oder falsch beurteilen kann.

Gleichwohl ist die Vereinbarung (und ggf. Abänderung) terminologischer Regeln in dem Sinne erfahrungsbedingt, als die sprachlichen Unterscheidungen, die wir mittels unserer Termini treffen, sich als mehr oder minder *zweckmäßig* erweisen können. Damit ein Lebewesen als Fisch bezeichnet wird, genügt es in der Zoologie z.B. nicht mehr, dass es im Wasser herumschwimmt etc., sondern es soll aufgrund der oben angedeuteten Regel noch hinzukommen, dass es sich in einer bestimmten Weise (und zwar anders als die Säugetiere) fortpflanzt. So können Prädikatenregeln die Unterscheidung von Gegenständen (in unserem Beispiel: zum Zwecke einer zoologischen Systematik) nicht nur verändern, sondern auch verschärfen.

Um Begriffsverwirrung zu vermeiden, ist es bei der expliziten Vereinbarung einer wissenschaftlichen Terminologie allerdings geboten, möglichst nahe an dem üblichen Sprachgebrauch zu

bleiben, d.h. den faktischen Sprachgebrauch von Fachwissenschaftlern möglichst treffend wiederzugeben, und dort, wo davon abgewichen wird, dies ausdrücklich zu begründen.

Dieses Problem verschärft sich in der Psychologie noch dadurch, dass manche psychologischen Fachtermini (z.B. Aggression) nicht nur von Psychologen sehr uneinheitlich gebraucht werden, sondern auch Eingang in die Umgangssprache gefunden haben. Die Psychologie unterscheidet sich aber von anderen Erfahrungswissenschaften (wie z.B. der Physik oder der Chemie) dadurch, dass sie mit ihrem Forschungsgegenstand – dem Menschen – sprechen kann und in vielen Praxisfeldern (z.B. Gesprächspsychotherapie, Beratung, Konfliktvermittlung) auf die Kommunikation mit ihm angewiesen ist.

1.3.3 Exkurs: Der Streit um den Aggressionsbegriff

Alltagssprachgebrauch

Im Alltagssprachgebrauch ist die Bezeichnung eines Verhaltens als Aggression mit einer Abwertung oder Verurteilung dieses Verhaltens verbunden: Der Aggressor ist der Angreifer, der sich damit ins Unrecht setzt. Die anderen werden zu den Angegriffenen, die sich verteidigen, quasi Notwehr begehen, deren Verhalten dadurch gerechtfertigt wird.

Etymologische Ursprünge

Etymologisch gesehen leitet sich das Wort ‚Aggression‘ von lat. *ag-gredior* her. Das bedeutet laut Stowasser: (1) herangehen; (2) sich an jemanden wenden, ihn angehen (*legatos*: zu gewinnen suchen; *pecunia*: zu bestechen suchen) – jedenfalls ein sehr zielstrebiges Herangehen an jemanden, das erst in einer Nebenbedeutung (wenn die friedlichen Mittel der Werbung und Bestechung versagen) angreifen bedeutet wie in unserem verkürzten Alltagssprachgebrauch; (3) unternehmen, beginnen, versuchen, an etwas gehen.

In diesem ursprünglichen Sinn des Wortes bezeichnet Aggression also zunächst nicht mehr und nicht weniger als tatkräftiges Handeln, zielstrebiges Verhalten, das Widerstände aus dem Weg räumt, ggf. auch Feinde: Erst in Verbindung mit Worten wie *hostes* (Feinde) wird die Tatkraft zum Angriff, nimmt das Wort ‚ag-gredior‘ die Bedeutung ‚angreifen‘ an.

Betrachten wir Aggression in dieser ursprünglichen und umfassenden Wortbedeutung, so folgt, dass Aggression nicht von vorn-

herein negativ zu bewerten ist (wie wir dies im Alltagssprachgebrauch gerne tun). Aggression ist nur "das sogenannte Böse" (Lorenz, 1963), während die Fähigkeit zur Aggression tatsächlich eine evolutionsbiologisch höchst funktionale Verhaltenskompetenz darstellt, ohne die weder individuelles Überleben noch Arterhaltung möglich wären.

Aus evolutionsbiologischer Sicht wird Aggression als eine Form des Konkurrenzverhaltens um fitnessbegrenzende Ressourcen (neben anderen) verstanden und nach Markl (1982) als ein Verhalten definiert, das geeignet und darauf gerichtet ist, die Fitness eines Konkurrenten zu mindern, indem ihm ein fitnessbegrenzendes Gut (z.B. Nahrung, Sexualpartner etc.) weggenommen oder vorenthalten wird, das stattdessen der Steigerung der Fitness des Aggressors zugute kommt.

Aggression als
Gegenstand der
Biologie

Trotz aller methodischer Ungenauigkeit besteht die bahnbrechende Leistung von Konrad Lorenz laut Bischof (1991) u.a. darin, die Aggression als *intraspezifisch* definiert und damit als eine prototypisch *soziale*, auf den Artgenossen gerichtete Verhaltensbereitschaft erkannt zu haben, die (entgegen Freuds Lehre vom Todestrieb) ursprünglich keine destruktive Pathologie darstellt, sondern im Bauplan des organismischen Verhaltens eine Funktion erfüllt.

Zusammen mit der Aggression sind auch aggressionshemmende Mechanismen stammesgeschichtlich evolviert, deren balancierende Wirkung erst dann versagt, wenn ein Lebewesen aus seiner natürlichen Ökologie fällt. Genau darin sieht Lorenz das existentielle Problem des Menschen, dessen Waffentechnik sich weit schneller entwickelt hat als die Hemmungsmechanismen, welche den Einsatz der Waffen unterbinden könnten.

In der Psychologie findet sich ein ähnlich umfassender Aggressionsbegriff teilweise in der Psychoanalyse, wenn sie davon ausgeht, dass jeder Kulturleistung ein gehöriges Maß an Aggression zugrunde liegt, aber auch darüber hinaus in so gut wie allen psychotherapeutischen Schulen, die in der Unterdrückung von Aggression z.B. eine Ursache für psychische und psychosomatische Erkrankungen sehen.

Aggressionsbegriffe
der Psychologie

In der psychologischen Aggressionsforschung haben dagegen bis in die 70er Jahre hinein die negativen Akzentuierungen des Aggressionsbegriffs dominiert. Dies gilt vor allem für den Aggressi-

Der behavioristische
Aggressionsbegriff

onsbegriff des Behaviorismus, für den ein Verhalten durch die schädlichen Auswirkungen, welche es für einen anderen Organismus hat, zur Aggression wird. So definiert Buss (1961): "Aggression is a behavior, which delivers noxious stimuli to another organism".

Schon bei anderen behavioristischen Autoren (vgl. Berkowitz, 1969) findet sich aber bereits eine Unterscheidung zwischen instrumenteller und feindseliger Aggression.

Der Versuch, diese Unterscheidung terminologisch streng zu fassen, löste in den 70er Jahren eine umfassende Diskussion des Aggressionsbegriffes aus, die am Ende zu einer Abkehr vom behavioristischen Methodenideal der Beschränkung auf das beobachtbare Verhalten (als einzig legitimen Gegenstand der Psychologie) geführt hat.

Intentionalität des Aggressionsbegriffs

War der Behaviorismus davon ausgegangen, dass Intentionen "private events" seien, die einer wissenschaftlichen Behandlung nicht zugänglich sind (Buss, 1961), so setzte sich nun die Auffassung durch, dass der Begriff der Aggression ohne Berücksichtigung der Intentionen, denen ein Verhalten dient, nicht definierbar ist (Werbik, 1971).

Da die Unterscheidung zwischen einer Aggression und einem Unfall (oder einer nur zufällig eingetretenen Schädigung eines anderen Organismus) auf Basis des Buss'schen Aggressionsbegriffs nicht möglich ist, definierte man Aggression nunmehr als eine Handlung, welche die Schädigung einer anderen Person beabsichtigt oder in Kauf nimmt.

Instrumentelle und feindselige Aggression

Die o.a. Unterscheidung zwischen instrumenteller und feindseliger Aggression lässt sich dann so fassen:

- *Instrumentelle Aggression* ist eine Handlung, welche die Schädigung einer anderen Person als Mittel zur Durchsetzung eigener Ziele einsetzt oder zumindest in Kauf nimmt.
- *Feindselige Aggression* ist eine Handlung, deren (eigentliches) Ziel es ist, eine andere Person zu schädigen.

Diese Unterscheidung hat eine Entsprechung im Strafrecht, wo zwischen den subjektiven Tatbestandsmerkmalen direkt/bedingt vorsätzlich und bewusst/unbewusst fahrlässig unterschieden wird (vgl. Groeben, Nüse & Gauler, 1992).

- Ein *direkter Vorsatz* besteht, wenn die objektiven Tatbe-

standsmerkmale "wissentlich und willentlich" herbeigeführt wurden.

- Ein *bedingter Vorsatz* besteht dagegen bereits dann, wenn der Eintritt des objektiven Tatbestandsmerkmals als Nebeneffekt einer Handlung "billigend in Kauf genommen wird".
- *Bewusste Fahrlässigkeit* liegt vor, wenn man von der Möglichkeit des Eintreffens des Tatbestandsmerkmals weiß, aber das Risiko in Kauf nimmt.
- *Unbewusste Fahrlässigkeit* liegt vor, wenn einem das Risiko nicht bekannt ist.

Wendet man diese Begrifflichkeit auf den Terminus ‚Aggression‘ an, so erweisen sich Handlungen, als deren Wirkung die Schädigung einer anderen Person eintritt, nur dann als nicht aggressiv, wenn die Schädigung unbewusst fahrlässig herbeigeführt wurde, während feindselige Aggressionen stets einen direkten Vorsatz zur Schädigung des anderen beinhalten und instrumentelle Aggressionen entweder direkten Vorsatz enthalten können (z.B. Schädigung des anderen, um ihn zum Nachgeben zu zwingen) oder einen bedingten Vorsatz (z.B. Aneignung durch Diebstahl). Ob Schädigungen, die durch bewusste Fahrlässigkeit herbeigeführt werden, als aggressiv gelten sollen, bleibt dagegen kontrovers. Dies ist nicht unproblematisch für die Dynamik von Konflikten, weil hier eine Perspektivendivergenz zwischen Täter und Opfer besteht.

Durch seine Definition über den Schädigungsaspekt hat der bis hier verwendete Aggressionsbegriff aber immer noch eine deutlich negative Konnotation und ist damit sowohl von seinen etymologischen Ursprüngen als auch vom Aggressionsbegriff der Verhaltensbiologie entfernt. Schädigung

Die Versuche zur terminologischen Klärung des Aggressionsbegriffs in den 70er Jahren haben auch diesbezüglich eine Rückbesinnung mit sich gebracht. Diese hatte zunächst – rein methodisch – die Schwierigkeiten zum Ausgangspunkt, welche eine objektive Definition des Begriffs der Schädigung mit sich bringt: So hatte Werbik (1971) noch gemeint, den Begriff der Schädigung durch eine Aufzählung verschiedener Arten von (schädigenden) Handlungsfolgen (zwar nicht definieren, aber) ersetzen zu können: Tötung, (körperliche) Verletzung usw.

Auch dabei zeigte sich jedoch die Tücke im Detail: Ist z.B. Sterbehilfe eine Aggression? Oder das Durchstechen der Nasenflügel?

Offensichtlich ist das nicht der Fall, wenn es der Wille der betroffenen Person ist (die sich Löcher in die Nasenflügel machen lässt, um darin Diamanten – oder anderen Schmuck – zu befestigen).

Konflikttheoretischer Aggressionsbegriff

Da also selbst Tötung und körperliche Verletzung nicht per se als Schädigung gelten können, hat das Ringen um den Gegenstand der Aggressionsforschung schließlich dazu geführt, den Schädigungsaspekt hintanzustellen und den Verstoß gegen den (vermuteten) Willen einer anderen Person zum entscheidenden Definitionsmerkmal der Aggression zu machen (Kempf, 1978; Werbik 1982). Damit ist der Aggressionsbegriff wieder in die Nähe seiner etymologischen Ursprünge zurückgekehrt.¹

Der konflikttheoretische Aggressionsbegriff

Eine Handlung gilt dann als aggressiv, wenn der Akteur meint, damit gegen den Willen einer anderen Person/anderer Personen zu verstoßen.

Weiterbestehen von Begriffsverwirrung

Ob man hierfür tatsächlich das Wort ‚Aggression‘ verwenden oder auf den Aggressionsbegriff ganz verzichten und stattdessen von Durchsetzungshandlungen (Kambartel, 1981) sprechen will, sei dahingestellt. Für letzteres spricht, dass die Diskussion um den Aggressionsbegriff in der Psychologie Mitte der 80er Jahre eingeschlafen ist. In den Friedenswissenschaften ist die o.g. Definition zwar mehr oder minder generell akzeptiert, in der Psychologie insgesamt jedoch besteht nach wie vor Begriffsverwirrung. Deshalb ist Vorsicht geboten und bei jeder Untersuchung genau darauf zu achten, wie der jeweilige Autor den Terminus ‚Aggression‘ verwendet und/oder wie er Aggression in seiner Untersuchung operationalisiert. Es steht zu vermuten, dass zumindest manche (scheinbar) widersprüchliche Ergebnisse der Aggressionsforschung de facto der Tatsache geschuldet sind, dass mit Aggression schlichtweg verschiedene Gegenstände bezeichnet werden.

Gemeinsamkeiten und Unterschiede der verschiedenen Aggressionsbegriffe

Zusammenfassend lassen sich in der Psychologie (mindestens) drei verschiedene Aggressionsbegriffe unterscheiden, die es sorgfältig auseinander zu halten gilt:

- Behavioristischer Aggressionsbegriff = Verletzendes Verhalten

1. So ist etwa Sterbehilfe, die auf Wunsch des Sterbenden erfolgt, keine Aggression, während es einiger Aggressivität – im Sinne der Durchsetzung eigener Ziele gegen den Willen des anderen – bedarf, sich der Bitte eines Sterbenden zu widersetzen, ihm Sterbehilfe zu leisten.

- *Biologischer Aggressionsbegriff* = Steigerung der eigenen Fitness auf Kosten der Fitness eines Konkurrenten
- *Konflikttheoretischer Aggressionsbegriff* = Durchsetzungshandlung

Während der behavioristische Aggressionsbegriff die Aggression von ihren (negativen) *Folgen* her definiert, verstehen der biologische und der konflikttheoretische Aggressionsbegriff die Aggression von ihrer *Funktion* bzw. ihren *Zielen* her und definieren die Aggression als eine bestimmte Form *organismischen Verhaltens* bzw. *menschlichen Handelns* in *Konkurrenz-* bzw. *Konfliktsituationen*.

Die hier getroffene Entscheidung zugunsten des konflikttheoretischen Aggressionsbegriffs ist Zweckmäßigkeitsüberlegungen geschuldet:

- gegenüber dem behavioristischen Aggressionsbegriff: wegen der im Behaviorismus bestehenden Ununterscheidbarkeit von Aggressionen und Unfällen;
- gegenüber dem biologischen Aggressionsbegriff: wegen der im Vergleich zu organismischem Verhalten weit höheren Flexibilität menschlichen Handelns, das nicht nur (objektseitig beschreibbare) biologische Funktionen hat, sondern auch (subjektseitig definierte) individuelle Ziele verfolgen kann.

Auf die damit angesprochene terminologische Unterscheidung zwischen funktionalem Verhalten und intentionalem Handeln wird noch zurückzukommen sein. Da es sich in beiden Fällen jedoch um theoriesprachliche Beschreibungsprädikatoren handelt, können wir ihre Verwendungsweise erst weiter unten und mit Bezug auf die ihnen zugrundeliegenden Erklärungsmodelle präziser herausarbeiten.

Die Bevorzugung des konflikttheoretischen Aggressionsbegriffs gegenüber dem des Behaviorismus erweist sich nicht nur abstrakt methodisch als zweckmäßig, sondern ganz konkret im Hinblick auf die Aufgaben der Konfliktforschung, die eine Unterscheidung zwischen subjektseitig definierten Durchsetzungshandlungen (Aggression) und objektseitig definiertem verletzendem Verhalten (Gewalt) als geboten erscheinen lassen: Wie wir sehen werden, ist es gerade diese Diskrepanz zwischen dem *äußeren Aspekt* von Handlungen (dem objektseitig definierten Verhalten und seinen Folgen) und ihrem *inneren Aspekt* (den subjektseitig definierten

Entscheidung zugunsten des konflikttheoretischen Aggressionsbegriffs

Innerer und äußerer Aspekt einer Handlung

Intentionen), welche die spezifische Dynamik von Konfliktverläufen zu erklären vermag.

Autoaggression

Aus der Begründung des konflikttheoretischen Aggressionsbegriffs relativ zu den Aufgaben der Konfliktforschung ergibt sich andererseits, dass diese Definition keineswegs auf alles anzuwenden wäre, wofür man in der Psychologie oder im Alltagsleben das Wort ‚Aggression‘ verwendet. So wäre es z. B. ziemlich unsinnig, *Autoaggression* als eine Durchsetzungshandlung zu definieren, bei der man seine eigenen Ziele gegen seinen eigenen Willen durchsetzt.

Der Begriff der Autoaggression, wie er in der klinischen Psychologie verwendet wird, ist vielmehr ein objektseitig definierter Begriff, der die Schädigung der eigenen Fitness beschreibt.

In der Psychoanalyse wird dieser faktischen Schädigung dann noch eine unbewusste Schädigungsabsicht unterstellt. Ob und wie weit die Unterstellung unbewusster Motive methodisch gerechtfertigt werden kann, werden wir weiter unten noch zu diskutieren haben.

1.3.4 Die logischen Partikel

Junktoren

Schon die Gebrauchssprache kennt eine Vielzahl an logischen Partikeln, mittels derer Elementaraussagen zu komplexeren Aussagen verknüpft werden können: Worte wie ‚und‘, ‚oder‘, ‚wenn-dann‘ usw.

Solche logischen Partikel, mittels derer je zwei Aussagen verknüpft werden, bezeichnet man als Junktoren. Die wichtigsten sind die Konjunktion, für die man in der Logik das Zeichen \wedge verwendet (dem ‚und‘ der natürlichen Sprache entsprechend), die Adjunktion \vee (einschließendes ‚oder‘), die Subjunktion \rightarrow (‚wenn-dann‘) und die Negation \neg (‚nicht‘).¹

Wertdefinite Aussagen

Was diese Verknüpfungen bedeuten, kann man festlegen, indem man den Wahrheitswert der zusammengesetzten Aussage als Funktion der Wahrheitswerte der miteinander verknüpften Teilaussagen darstellt. Die Teilaussagen können entweder elementar

1. Die Negation wird zu den Junktoren gerechnet, da sie ebenfalls als Wahrheitsfunktion (siehe unten) auffassbar ist, obwohl sie nicht zwei Aussagen miteinander verknüpft.

oder auch ihrerseits bereits logisch zusammengesetzt sein. Am gebräuchlichsten ist die von G. Frege (1879, zit. n. Lorenz, 1984) begründete klassische Logik,¹ die von wertdefiniten Aussagen ausgeht, also voraussetzt, dass jede Aussage entweder wahr (w) oder falsch (f) ist (tertium non datur).

Die Wahrheitsfunktion der Junktoren wird gewöhnlich mit Hilfe von Wahrheitstafeln dargestellt:

Negation: ‚nicht‘

Die Negation

A	$\neg A$
w	f
f	w

Die Negation einer Aussage ist genau dann wahr, wenn die Aussage falsch ist (und umgekehrt).

Konjunktion: ‚und‘

Die Konjunktion

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

Die Konjunktion zweier Aussagen ist genau dann wahr, wenn beide Aussagen wahr sind.

Adjunktion: (einschließendes) ‚oder‘

Die Adjunktion

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

Die Adjunktion zweier Aussagen ist genau dann wahr, wenn mindestens eine der beiden Aussagen wahr ist.

Subjunktion: ‚wenn - dann‘

Die Subjunktion

A	B	$A \rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Die Subjunktion ist genau dann falsch, wenn der Vordersatz (A) wahr und der Nachsatz (B) falsch ist.

1. Der Beweis der Vollständigkeit des klassischen Logikkalküls wurde allerdings erst 1930 von K. Gödel erbracht.

Logisch wahre Sätze Im vorangegangenen Abschnitt hatten wir gefordert, dass die Prädikatorenregeln und Definitionen logisch korrekt, d.h. widerspruchsfrei sein müssen. Diese Forderung nach Widerspruchsfreiheit (nicht nur der terminologischen Regeln, sondern auch der Theorien) hat einen guten Grund. Aus widersprüchlichen Voraussetzungen (Prämissen) kann nämlich jede beliebige Behauptung gefolgert werden; und zwar allein aufgrund des logischen Widerspruches der Form $[A \wedge \neg A]$ und unabhängig davon, ob die daraus gezogene Schlussfolgerung, die sogenannte Conclusio (B), wahr oder falsch ist.

Wir können uns davon leicht überzeugen, indem wir die Wahrheitstafel der Subjunktion $[A \wedge \neg A] \rightarrow B$ erstellen.

A	B	$\neg A$	$A \wedge \neg A$	$[A \wedge \neg A] \rightarrow B$
w	w	f	f	w
w	f	f	f	w
f	w	w	f	w
f	f	w	f	w

Wie wir sehen, ist die Subjunktion $[A \wedge \neg A] \rightarrow B$ stets wahr, und zwar unabhängig davon, ob die Teilaussagen A und B wahr sind, ja sogar unabhängig davon, was diese Teilaussagen überhaupt bedeuten. Eine solche Aussage, die allein aufgrund ihrer logischen Form wahr ist, bezeichnet man als Tautologie, oder, wie wir auch sagen können, als (formal) logisch wahr.

Logische
Schlussverfahren

In den Erfahrungswissenschaften sind Tautologien insofern unbrauchbar, als sie keinerlei Informationsgehalt bezüglich der darin vorkommenden Teilaussagen haben. In der Logik, die zwar nicht Teil der Erfahrungswissenschaften ist, aber zu dem methodischen Fundament gehört, auf welches die Erfahrungswissenschaften aufbauen, spielen logisch wahre Sätze dagegen insofern eine bedeutsame Rolle, als sie die Grundlage der logischen Schlussverfahren bilden (ohne die auch keine Erfahrungswissenschaft auskommt). Von besonderer Bedeutung für die Erfahrungswissenschaften sind dabei zwei Schlussverfahren, die man als modus ponens und modus tollens bezeichnet.

Der modus ponens

Der modus ponens schließt aus den Prämissen A und $A \rightarrow B$ auf die Wahrheit der Aussage B. Formal:

$$(A \wedge [A \rightarrow B]) \rightarrow B$$

Wie man sich mittels der entsprechenden Wahrheitstafel leicht überzeugen kann, ist diese Aussage logisch wahr. Man sagt daher auch: $(A \wedge [A \rightarrow B])$ impliziert B.

Dasselbe gilt für den modus tollens, der aus den Prämissen $A \rightarrow B$ und $\neg B$ darauf schließt, dass die Aussage A falsch ist. Formal: Der modus tollens

$$([A \rightarrow B] \wedge \neg B) \rightarrow \neg A$$

Entsprechend können wir auch sagen: $([A \rightarrow B] \wedge \neg B)$ impliziert $\neg A$.

Neben den Junktoren verwendet man in der Logik Quantoren, mit deren Hilfe sich allgemeine Sätze der Form "Alle x sind P" bzw. "Es gibt (mindestens) ein x, das P ist" bilden lassen. Quantoren

Der Allquantor (\forall) stellt dabei gleichsam eine Verallgemeinerung der Konjunktion dar: Universelle Aussagen

$$\forall x. A(x) \Leftrightarrow A(x_1) \wedge A(x_2) \wedge A(x_3) \wedge \dots \text{ (usw.)}$$

Daraus folgt: Die universelle Aussage (Allaussage) $\forall x. A(x)$ (sprich: "für alle x gilt A(x)") ist genau dann wahr, wenn die Teilaussage A(x) für jedes einzelne x wahr ist. Sie ist dagegen falsch, sobald sich ein x findet, für das A(x) falsch ist.

Handelt es sich bei der Aussage A(x), deren universelle Geltung behauptet wird, um eine empirische Aussage (d.h. um eine Aussage, über deren Wahrheit oder Falschheit mittels Beobachtung zu entscheiden ist) und ist der Gegenstandsbereich (= die Menge aller x) unendlich, so kann die Allaussage $\forall x. A(x)$ offensichtlich nicht definitiv bewiesen werden. Denn dazu müsste man die Aussage A(x) an unendlich vielen Fällen empirisch bewahrheiten und würde nie zu einem Ende kommen. Die Allaussage $\forall x. A(x)$ kann jedoch widerlegt (falsifiziert) werden, und sie ist widerlegt, sobald sich ein Gegenbeispiel findet.

Ist A(x) dagegen eine sachlogische Behauptung, über deren Wahrheit oder Falschheit (wie z.B. über die Sätze der Mathematik) entschieden werden kann, ohne dass es dazu einer Beobachtung bedarf, so kann die universelle Aussage $\forall x. A(x)$ auch bei unendlichem Gegenstandsbereich sehr wohl bewiesen werden, und zwar dann, wenn es gelingt, eine Regel anzugeben, nach der die Aussage A(x) für jedes einzelne x bewiesen werden kann. Eine solche Regel ist z.B. das Prinzip der vollständigen Induktion in der Mathematik:

Die vollständige
Induktion

Das Prinzip der vollständigen Induktion

Behauptung: Für alle natürlichen Zahlen $n = 1, 2, 3, \dots$ gilt $A(n)$

Beweis:

(1)	$A(1)$
(2)	$A(n) \rightarrow A(n+1)$

Also gilt $A(n)$ auch für $n=2$, also auch für $n=3$, usw.

Partikuläre Aussagen

Genau das umgekehrte Problem besteht für die Überprüfung von partikulären Aussagen der Form "Es gibt (mindestens) ein x , für das $A(x)$ zutrifft": Der Einsquantor (\exists) ist als Verallgemeinerung der Adjunktion definiert:

$$\exists x. A(x) \Leftrightarrow A(x_1) \vee A(x_2) \vee A(x_3) \vee \dots \text{ (usw.)}$$

Daraus folgt: Die partikuläre Aussage (Existenzaussage)

$\exists x. A(x)$ ist genau dann falsch, wenn die Teilaussage $A(x)$ für jedes einzelne x falsch ist. Sie ist dagegen wahr, sobald sich ein x findet, für das $A(x)$ zutrifft.

Handelt es sich bei der Aussage $A(x)$, deren Existenz behauptet wird, um eine empirische Aussage, so folgt, dass sie bei unendlichem Gegenstandsbereich nicht definitiv widerlegt werden kann, da man dazu die Aussage $A(x)$ an unendlich vielen Fällen empirisch widerlegen müsste (also nie zu einem Ende kommen würde). Die Existenzaussage $\exists x. A(x)$ kann jedoch bewiesen werden, und sie ist bewiesen, sobald sich ein Beispiel findet, für das $A(x)$ zutrifft.

1.3.5 Wahrheit und Wirklichkeit

Ein definitorischer
Zirkel

Nach verbreitetem Vorverständnis, dessen philosophische Tradition bis auf Aristoteles zurückreicht, gilt eine Aussage dann als wahr, wenn sie mit der Wirklichkeit übereinstimmt. Beispiele sind die Erklärung Kants, dass "Wahrheit eine Übereinstimmung unserer Begriffe mit dem Objekte" ist, oder Hegels Erklärung, "dass die Idee nur *vermittelt* des Seyns, und umgekehrt das Seyn nur *vermittelt* der Idee, *das Wahre*" ergibt (zit. n. Lorenz, 1996).

Gegen eine solche Definition des Wahrheitsbegriffs spricht zumindest zweierlei: Zum einen müsste der Terminus ‚Wirklichkeit‘

bereits definiert sein – und zwar unabhängig vom Wahrheitsbegriff, weil wir uns andernfalls in einen definitorischen Zirkel begeben, d.h. Wahrheit letztlich durch sich selbst definieren. Zum anderen können immer nur Aussagen *miteinander* verglichen werden, nicht jedoch mit einer wie immer auch gearteten Wirklichkeit; es sei denn, dass diese in Form von Aussagen dargestellt ist. So können auch (empirische) Theorien nicht einfach mit Beobachtungen verglichen werden, sondern nur mit Aussagen über das Beobachtete.

Ein Ausweg aus diesem Dilemma ergibt sich, wenn wir uns darauf besinnen, dass wir mit der Behauptung einer Aussage eine gewisse Verpflichtung eingehen: nämlich die Verpflichtung, die Aussage auf Verlangen verteidigen zu können. Kambartel (1968a, 15) weist in diesem Zusammenhang darauf hin, dass es im Allgemeinen nicht angeht, *unüberlegt* Behauptungen aufzustellen. Das überlegte oder methodische Behaupten hat nämlich ein einleuchtendes praktisches Ziel. Es setzt uns in die Lage, Möglichkeiten und Grenzen des Handelns, deren *andere* sich versichert haben, für unser *eigenes* Handeln in Rechnung zu stellen. Da wir davon jederzeit Gebrauch machen, ist der Verlass auf die Behauptungen anderer ein unverzichtbares Element unserer eigenen Lebenspraxis geworden. D.h. es besteht ein begründetes Interesse daran, dass die Institution des Behauptens funktionsfähig bleibt, dass nur solche Aussagen behauptet werden sollen, die sich auf jeden Fall (und nicht nur gegenüber einem unkritischen oder sachkundigen Opponenten) verteidigen lassen.

Die Institution des Behauptens

Die Prädikation einer Aussage als wahr lässt sich im Anschluss an Frege als eine Betonung des Behauptungscharakters (Kambartel, 1968b, 102) rekonstruieren, als das ausdrückliche Versprechen, die mit der Behauptung der Aussage verbundenen Verteidigungspflichten auf Verlangen einzulösen. Die Prädikation einer Aussage als falsch lässt sich entsprechend als das ausdrückliche Versprechen rekonstruieren, die mit dem Bestreiten der Aussage verbundenen Verteidigungspflichten einzulösen.

Wahrheit

Die mit der Metaprädikation von Aussagen als wahr oder falsch getroffene Unterscheidung ist somit die zwischen solchen Aussagen, die wir für unser Handeln benutzen wollen, und solchen, die wir für unser Handeln ausdrücklich nicht benutzen wollen.

Wirklichkeit

Den Terminus ‚Wirklichkeit‘ können wir einführen, indem wir neben der Redeweise "Aussagen stellen Sachverhalte dar" auch noch die Redeweise "Wahre Aussagen stellen wirkliche Sachverhalte dar" zulassen.

Es wäre jedoch ein Missverständnis, wollte man übersehen, dass damit nur der Gebrauch einiger Termini (wahr, falsch, Sachverhalt, wirklich) geregelt wurde, und etwa zu glauben, die Wahrheit von Aussagen könne nicht mehr von der Erfahrung abhängen. Auf einem Missverständnis – allerdings anderer Art – beruht auch die in der Psychologie weit verbreitete Meinung, die Verifikation (theoretischer) Aussagen müsse in jedem Fall auf (empirischer) Erfahrung beruhen.

In der Philosophie und Wissenschaftstheorie gilt heute als gesichert, dass es verschiedene Arten von Aussagen gibt, die unterschiedliche Verteidigungspflichten nach sich ziehen und somit verschiedene Arten von Wahrheit/Falschheit definieren.

Vernunft- und
Tatsachenwahrheiten

Bereits Leibniz (1646–1716) unterschied *Verités de Raisonement*, also Vernunftwahrheiten, von *Verités de Fait*, also Tatsachenwahrheiten: "Vernunftwahrheiten sind notwendig und ihr Gegenteil ist unmöglich, Tatsachenwahrheiten sind kontingent und ihr Gegenteil ist möglich" (zit.n. Hoyningen-Huene, 1997, 230). Während Vernunftwahrheiten durch (logische) Analyse gefunden werden können, also rein verstandesmäßig zugänglich sind, reicht dieses Verfahren bei Tatsachenwahrheiten nicht aus. Bei kontingenten Sätzen kann man, so lange man nur den Verstand zu Rate zieht, nicht entscheiden, ob sie wahr oder falsch sind. Eben dies besagt ja die Kontingenz. Um zu entscheiden, ob eine kontingente Aussage wahr ist, benötigt man daher etwas, das außerhalb der Vernunft liegt, nämlich die Empirie.

Wie bereits Kant festgestellt hat, fließen in dieser von Leibniz getroffenen Unterscheidung jedoch zwei Dinge ineinander, die man besser getrennt halten sollte: Das ist zum einen die Frage nach dem Verhältnis der in einer Aussage verwendeten Termini zueinander; und zum anderen die Frage danach, worauf man sich bei der Verteidigung der Aussage zu berufen hat.

Analytische und
synthetische
Behauptungen

Behauptungen, die sich allein aufgrund der Terminologie (und der Logik) beweisen lassen (z.B. "Junggesellen sind unverheiratet"), bezeichnet Kant als *analytisch* und unterscheidet sie von

den *synthetischen* Behauptungen (z.B. "Der eben merkliche Unterschied zweier Stimuli ist proportional zum Standardreiz").

Behauptungen, zu deren Verteidigung man die Empirie (also Beobachtung und/oder Experiment) benötigt, bezeichnet Kant als *a posteriori* und unterscheidet sie von den Behauptungen *a priori*.

Behauptungen
a priori und
a posteriori

Kant hat diese beiden Unterscheidungen eingeführt, um sie miteinander zu kombinieren. Dabei ergeben sich zwei Kombinationen, die relativ unproblematisch sind.

Unproblematisch sind erstens die analytischen Behauptungen, die man *a priori* begründen kann. Natürlich braucht man keine Beobachtung, um z.B. zu beweisen, dass Junggesellen unverheiratet sind. Dazu genügt es, auf die Verwendungsregeln des Prädikators ‚Junggeselle‘ zu rekurren.

Ebenfalls unproblematisch sind die synthetischen Behauptungen, die man *a posteriori* begründen muss. Natürlich bedürfen wir der Beobachtung, um z.B. die Geltung des Weber'schen Gesetzes zu überprüfen.

	Analytisch	Synthetisch
A priori	Unproblematisch	Fraglich
A posteriori	Unsinnig	Unproblematisch

Tab. 1.3.1: Kombination der Unterscheidungen *analytisch vs. synthetisch* und *a priori vs. a posteriori*

Drittens ist auch die Kombination von analytisch und *a posteriori* in dem Sinne unproblematisch, als es keinen Sinn macht, Behauptungen empirisch nachprüfen zu wollen, deren Geltung sich bereits aus der Bedeutung der darin verwendeten Begriffe ergibt.

Problematisch ist diese Kombination allerdings insofern, als in der neueren Psychologie immer wieder der Verdacht aufgekommen ist, dass bestimmte, vermeintlich empirische Hypothesen tatsächlich analytisch zu begründen sind und somit bereits *a priori* gelten (z.B. Postman, 1947; Smedslund, 1976, 1991; Brandtstädter, 1982, 1984, 1987; Holzkamp, 1986). Sollte sich dieser Verdacht erhärten, so wäre manche empirische Untersuchung in der Psychologie bloße Pseudoempirie. Der Verdacht richtet sich dabei freilich nicht gegen die Geltung der fraglichen Hypothesen, sondern gegen die (unangemessene) Art und Weise, wie die Geltung dieser Hypothesen überprüft wird. Auch der Wert der Hypothesen – genauer: ihre Zulässigkeit innerhalb einer Erfahrungswis-

Pseudoempirie

senschaft – wird dadurch nicht in Frage gestellt. Trifft es aber zu, dass eine bestimmte Hypothese analytisch zu begründen ist, dann sagt sie (wie z.B. die Aussage "Junggesellen sind unverheiratet") allerdings nichts mehr über die empirische Wirklichkeit aus, wohl jedoch über die (sachlogische) Struktur des fraglichen Gegenstandsbereiches; d.h. es handelt sich dann nicht mehr um eine empirische, sondern um eine strukturelle Gesetzmäßigkeit.

Dass psychologische Theorien (wie übrigens die Theorien der harten Naturwissenschaften auch) nicht ausschließlich aus empirischen Aussagen bestehen, sondern auch a priorische Anteile enthalten, braucht uns dabei nicht weiter zu beunruhigen, ganz im Gegenteil. Wie wir noch sehen werden, versetzen uns die a priorischen Theorieanteile überhaupt erst in die Lage, Theorien zu formulieren, die in sich geschlossen sind und mehr als nur eine Aneinanderreihung von unverbunden nebeneinanderstehenden Hypothesen.

Synthetische
Behauptungen
a priori

Als fraglich erscheint die Möglichkeit synthetischer Behauptungen a priori. Kant war der Meinung, dass z.B. die Sätze der Arithmetik solche synthetischen Wahrheiten a priorischer Natur darstellen. Diese Auffassung machen sich auch Kamlah & Lorenzen (1967, 219ff) zu eigen, indem sie darauf hinweisen, dass die Sätze der Arithmetik synthetische Aussagen sind, wenngleich zwar nicht Aussagen über in der Natur vorfindliche Gegenstände, sondern über Gegenstände, die in der Arithmetik erst nach bestimmten Regeln konstruiert werden. Da die Verteidigung dieser Sätze es jedoch nicht erforderlich macht, auf die Wahrheit von Elementaraussagen zurückzugehen, sondern aufgrund der Konstruktionsregeln für die Gegenstände der Arithmetik erfolgen kann, sind sie a priori wahr.¹

Diese Auffassung, wonach es synthetische Wahrheiten a priori gibt, ist in der Erkenntnis- und Wissenschaftstheorie jedoch bis heute umstritten. Tatsächlich verschwindet der Unterschied zu den analytisch wahren Sätzen, wenn man die Mathematik als eine formale Sprache versteht und entsprechend die Konstruktionsregeln, nach denen die Gegenstände der Mathematik konstruiert sind, als terminologische Regeln.

1. In diesem Sinne werden wir unten sehen, dass das Gesetz der großen Zahlen, wonach die relative Häufigkeit zufälliger Ereignisse mit wachsender Zahl der Beobachtungen gegen die Wahrscheinlichkeit des Ereignisses strebt, sich unmittelbar aus den Konstruktionsprinzipien für Zufallsgeneratoren herleiten lässt (vgl. Lorenzen, 1978).

Ohne diese philosophische Diskussion weiter zu vertiefen, schließen wir uns im Folgenden einem in der Psychologie üblichen Sprachgebrauch an und fassen die analytischen Aussagen (im engeren Sinne) und die synthetischen Aussagen a priori unter dem Oberbegriff *sachlogisch* zusammen: Damit unterscheiden wir sie zugleich von den (formal) *logisch* wahren bzw. falschen Aussagen, die bereits aufgrund ihrer logischen Form (und ungeachtet der Bedeutung der darin enthaltenen Teilaussagen) wahr oder falsch sind, und von den *empirischen* Aussagen, zu deren Begründung (über die Regeln der Logik, der Terminologie und der Mathematik hinaus) eine Beobachtung (d.h. die Verteidigung einer Elementaraussage) erforderlich ist. Als sachlogisch werden somit all jene Aussagen bezeichnet, zu deren Verteidigung es (über die Regeln der Logik hinaus) der Regeln, der Terminologie und/oder der Mathematik bedarf, zu deren Beweis oder Widerlegung jedoch keine Beobachtung erforderlich ist.

Strukturelle und empirische Theorieanteile

Entsprechend unterscheiden wir zwischen *strukturellen* Theorieanteilen, die (sach)logisch zu begründen sind, und *empirischen* Theorieanteilen. Daraus ergibt sich die folgende Typologie des Wahrheitsbegriffs:

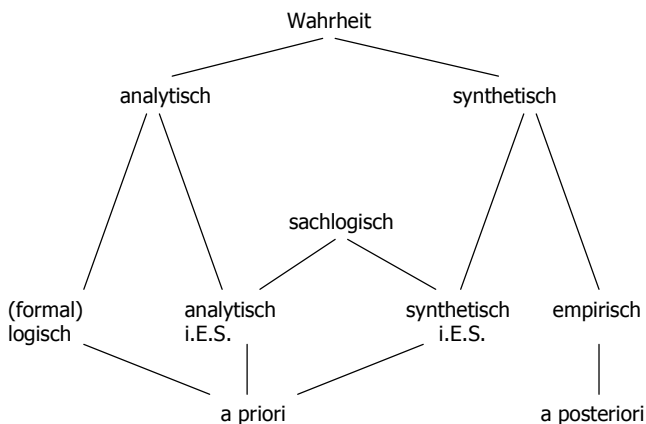


Abb. 1.3.2: Überblick über die verschiedenen Wahrheitsbegriffe

Kontrollfragen

- Was versteht man unter beschreibungssprachlichen Termini? Geben Sie Beispiele!
- Welche Arten von Sätzen lassen sich unterscheiden? Wie wird diese Unterscheidung eingeübt?

- Was ist eine Elementarsaussage, und welche Form hat sie?
- Worin besteht der Unterschied zwischen affirmativen und negativen Elementaraussagen?
- Worin besteht der Unterschied zwischen Nominatoren und Prädikato-
ren?
- Welche Arten von Nominatoren lassen sich unterscheiden?
- Worin besteht der Unterschied zwischen Beobachtungsprädikato-
ren, Deutungsprädikato- und Handlungsprädikato-
ren? Geben Sie ein Beispiel!
- Was versteht man unter einem theoretischen Konstrukt?
- Worin besteht die exemplarische Einführung von Prädikato-
ren?
- Welche Arten von Prädikato- können exemplarisch eingeübt wer-
den?
- Welche Arten von Prädikato- gehören zur Beschreibungssprache?
- Erläutern Sie, warum Handlungsprädikato- keine Beobachtungspräd-
ikato- sind! Geben Sie ein Beispiel!
- Geben Sie
 - a) eine beobachtungssprachliche,
 - b) eine theoriesprachliche
 Beschreibung von Pawlows Konditionierungsexperimenten!
- Erläutern Sie anhand von Pawlows Experimenten den Unterschied
zwischen Beobachtungssprache und Theoriesprache!
- Was versteht man unter der Operationalisierung theoriesprachlicher
Termini?
- Worin besteht die Kontrollphase klassischer Konditionierungsexperi-
mente, und wozu dient sie?
- Worin besteht der Unterschied zwischen einstelligen und mehrstelligen
Prädikato-? Geben Sie Beispiele!
- Was sind Relato-?
- Durch welche Eigenschaften sind Ordnungsrelato- charakterisiert?
- Welche Voraussetzungen müssen erfüllt sein, damit ein zweistelliger
Prädikato- eine Ordnungsrelation definiert?
- Welche Voraussetzung muss erfüllt sein, damit man eine Ordinalskala
(Rangskala) konstruieren kann?
- Wie kann man quantitative Begriffe (z.B. Länge) definieren?
- Was versteht man unter einer Prädikato-regel? Welche Form(en)
kann sie haben? Wozu dient sie?
- Wie kann die Geltungsprüfung von Elementaraussagen erfolgen,
 - a) wenn die darin vorkommenden Beschreibungsprädikato- ledig-
lich exemplarisch bestimmt sind,
 - b) wenn der Gebrauch der Prädikato- bereits durch Prädikato-
regeln abgesichert wurde?
- Durch welche Prädikato-regeln ist die Inhaltsgleichheit zweier Präd-
ikato- definiert?
- Was bedeutet die Aussage, dass zwei Prädikato-
 - a) denselben Sachverhalt darstellen,
 - b) dieselbe Bedeutung haben?
- Worin unterscheiden sich inhaltsgleiche Prädikato-?
- Welche Formen der Unverträglichkeit von Prädikato- gibt es?
- Durch welche Prädikato-regel(n) ist der konträre Gegensatz defi-

niert?

- Durch welche Prädikatorenregel(n) ist der kontradiktorische Gegensatz definiert?
- Welcher der beiden Gegensätze ist schärfer?
- Welche der folgenden Prädikatorenregeln treffen zu?
 - a) Die Prädikatoren P und Q sind unverträglich \Rightarrow P und Q sind konträre Gegensätze.
 - b) Die Prädikatoren P und Q sind unverträglich \Rightarrow P und Q widersprechen einander.
 - c) Die Prädikatoren P und Q sind unverträglich \Rightarrow P und Q sind kontradiktorische Gegensätze.
 - d) Die Prädikatoren P und Q sind unverträglich \Rightarrow P ist das Gegenteil von Q.
 - e) Die Prädikatoren P und Q sind konträre Gegensätze \Rightarrow P und Q sind unverträglich.
 - f) Die Prädikatoren P und Q sind konträre Gegensätze \Rightarrow P und Q widersprechen einander.
 - g) Die Prädikatoren P und Q sind konträre Gegensätze \Rightarrow P und Q sind kontradiktorische Gegensätze.
 - h) Die Prädikatoren P und Q sind konträre Gegensätze \Rightarrow P ist das Gegenteil von Q.
 - i) P ist das Gegenteil von Q \Rightarrow P und Q sind unverträglich.
 - j) P ist das Gegenteil von Q \Rightarrow P und Q widersprechen einander.
 - k) P ist das Gegenteil von Q \Rightarrow P und Q sind konträre Gegensätze.
 - l) P ist das Gegenteil von Q \Rightarrow P und Q sind kontradiktorische Gegensätze.
- Wie können polar-konträre Prädikatoren eingeführt werden?
- Durch welche Prädikatorenregel sind die Ordnungsrelatoren ‚länger als‘ und ‚kürzer als‘ miteinander verbunden?
- Worin besteht der Unterschied zwischen Prädikatorenregeln und Definitionen?
- Wie lautet die Definition der verzögerten Konditionierung?
- Welche logische Form haben Definitionen?
- Worin besteht der Unterschied zwischen Definiens und Definiendum?
- Was bedeutet die Aussage, dass Definiens und Definiendum inhaltsgleich sind?
- Worin besteht der Unterschied zwischen Sprachregeln (Prädikatorenregeln und Definitionen) und empirischen Hypothesen?
- Worauf muss man bei der Vereinbarung von Sprachregeln achten?
- Was versteht man alltagssprachlich unter Aggression?
- Wie leitet sich der Terminus ‚Aggression‘ etymologisch her?
- Wie ist der Terminus ‚Aggression‘ in der Biologie definiert?
- Wie ist der Terminus ‚Aggression‘ im Behaviorismus definiert?
- Warum wurde der behavioristische Aggressionsbegriff später fallen gelassen?
- Worin besteht der Unterschied zwischen instrumenteller und feindlicher Aggression?
- Worin besteht der Unterschied zwischen einem direkten und einem bedingten Vorsatz?
- Worin besteht der Unterschied zwischen bewusster und unbewusster

Fahrlässigkeit?

- Welche Beziehungen bestehen zwischen instrumenteller vs. feindseliger Aggression einerseits und (direktem bzw. bedingtem) Vorsatz vs. (bewusster bzw. unbewusster) Fahrlässigkeit andererseits?
- Wie lautet der konflikttheoretische Aggressionsbegriff?
- Worin bestehen Gemeinsamkeiten und Unterschiede des behavioristischen, des biologischen und des konflikttheoretischen Aggressionsbegriffs?
- Welchen Zweckmäßigkeitserwägungen lassen sich zugunsten des konflikttheoretischen Aggressionsbegriffs vorbringen?
- Was versteht man unter dem äußeren und dem inneren Aspekt einer Handlung?
- Wie ist der Begriff der Autoaggression definiert?
- Wie heißen die folgenden logischen Verknüpfungen, durch welches Symbol werden sie dargestellt und wie lautet die jeweilige Wahrheitstafel, durch welche sie definiert sind?
 - a) "nicht",
 - b) "und",
 - c) (einschließendes) "oder",
 - d) "wenn-dann".
- Was versteht man unter wertdefiniten Aussagen?
- Ergänzen Sie die folgenden Sätze!
 - a) Die Aussage $\neg A$ ist genau dann wahr, wenn ...
 - b) Die Aussage $A \wedge B$ ist genau dann wahr, wenn ...
 - c) Die Aussage $A \vee B$ ist genau dann wahr, wenn ...
 - d) Die Aussage $A \rightarrow B$ ist genau dann falsch, wenn ...
- Warum muss man auf die Widerspruchsfreiheit von Theorien achten?
- Zeigen Sie mittels der Wahrheitstafel, dass die Subjunktion $[A \wedge \neg A] \rightarrow B$ stets wahr ist!
- Was versteht man unter einer Tautologie?
- Inwiefern sind Tautologien in den Erfahrungswissenschaften
 - a) unbrauchbar,
 - b) unverzichtbar?
- Überprüfen Sie mittels der Wahrheitstafeln, ob es sich bei den folgenden Aussagen um logisch wahre Sätze handelt:
 - a) $A \vee [A \wedge B]$
 - b) $A \vee [A \rightarrow B]$
 - c) $A \wedge [\neg A]$
- Wie lautet die logische Schlussform
 - a) des modus ponens,
 - b) des modus tollens?
- Zeigen Sie mittels der Wahrheitstafeln, dass es sich dabei um Tautologien handelt!
- Was versteht man unter der Falsifikation einer Aussage?
- Wie ist der Allquantor definiert?
- Was versteht man unter einer universellen Aussage? Wann ist sie wahr? Wann ist sie falsch?
- Unter welchen Voraussetzungen können universelle Aussagen verifiziert (= bewiesen) werden, und unter welchen Voraussetzungen ist lediglich ihre Falsifikation (= Widerlegung) möglich?

- Warum können empirische Allaussagen mit unendlichem Gegenstandsbereich nicht definitiv bewiesen werden?
- Wodurch können empirische Allaussagen mit unendlichem Gegenstandsbereich falsifiziert werden?
- Wie können Aussagen der Form $\forall x. A(x)$ bewiesen werden, wenn es sich bei der Teilaussage $A(x)$ um eine sachlogische Behauptung handelt?
- Worin besteht das Prinzip der vollständigen Induktion?
- Wie ist der Einsquantor definiert?
- Was versteht man unter einer partikulären Aussage? Wann ist sie wahr? Wann ist sie falsch?
- Unter welchen Voraussetzungen können partikuläre Aussagen falsifiziert werden, und unter welchen Voraussetzungen ist lediglich ihre Verifikation möglich?
- Warum können empirische Existenzaussagen mit unendlichem Gegenstandsbereich nicht definitiv widerlegt werden?
- Wodurch können empirische Existenzaussagen mit unendlichem Gegenstandsbereich verifiziert werden?
- Kritisieren Sie die Definition "Eine Aussage ist wahr, wenn sie mit der Wirklichkeit übereinstimmt"!
- Welchem (praktischen) Interesse dient die Unterscheidung von wahren und falschen Aussagen?
- Wie lässt sich die (Meta-)Prädikation einer Aussage als wahr oder falsch rekonstruieren?
- Was bedeuten die Sätze:
 - a) "Die Aussage A ist wahr" und
 - b) "Die Aussage A ist falsch"?
- Wie lässt sich der Terminus ‚Wirklichkeit‘ rekonstruieren?
- Wie unterscheidet Leibniz zwischen Vernunftwahrheiten und Tatsachenwahrheiten?
- Wie unterscheidet Kant
 - a) zwischen analytischen und synthetischen Wahrheiten und
 - b) zwischen Wahrheiten a priori und Wahrheiten a posteriori?
- Welche der möglichen Kombinationen von analytisch vs. synthetisch mit a priori vs. a posteriori sind
 - a) unproblematisch,
 - b) unsinnig,
 - c) fraglich,und warum ist das so?
- Was versteht man unter Pseudoempirie? Geben Sie ein Beispiel!
- Worin besteht der Unterschied zwischen
 - a) (formal) logischen,
 - b) sachlogischen und
 - c) empirischen Aussagen?
- Worin besteht der Unterschied zwischen
 - a) strukturellen und
 - b) empirischen Theorieanteilen?

1.4

Aufgabenverständnis, Gegenstandsverständnis und Wissensideal der Psychologie

Psychologie und
Nachbardisziplinen

Jeder Wissenschaft geht eine Entscheidung darüber voraus, was als wissenschaftlich erachtet wird. Dass diese Entscheidung in der Psychologie nicht gerade einheitlich ausgefallen ist, hängt mit der Komplexität ihres Gegenstandsbereiches zusammen. Psychologie als Wissenschaft von der subjektiven Welt des Menschen umfasst ein äußerst breites Wissensgebiet, das nicht nur die Funktionsweise des bewussten Erlebens, sondern auch dessen Ursachen und Konsequenzen umfasst. Damit gehören biogenetische und physiologische Ursachen menschlichen Erlebens und Handelns ebenso zum Gegenstand der Psychologie wie unbewusste Informationsverarbeitungsprozesse und Prozesse der sozialen Konstruktion der Wirklichkeit, das daraus resultierende menschliche Verhalten und Handeln und schließlich auch dessen Produkte, welche die Lebenswelt des Menschen verändern und darüber wiederum eine Rückkoppelung auf das individuelle Erleben und Handeln haben. Entsprechend gliedert sich die Psychologie heute in eine Vielzahl von Teilgebieten, von denen manche einen fließenden Übergang zu Nachbardisziplinen haben: So geht die Physiologische Psychologie beinahe nahtlos in die Physiologie über, die Gegenstand der Medizin und der Biologie ist. Dasselbe gilt auch für die Tierpsychologie, die zugleich Überschneidungen mit der Verhaltensbiologie hat, sowie für Teile der Klinischen Psychologie, die wiederum der Psychiatrie eng verwandt ist. Andere Teilgebiete, wie z.B. Sozialpsychologie, Arbeits-, Betriebs- und Organisationspsychologie, Medienpsychologie und Friedenspsychologie, überschneiden sich teilweise mit Sozialwissenschaften, wie z.B. mit der Soziologie, mit den Kommunikationswissenschaften und der Politologie. Wiederum andere, wie z.B. die Rechtspsychologie oder die Kulturpsychologie, haben enge Berührungspunkte mit den entsprechenden Geisteswissenschaften.

Angesichts dieser Heterogenität der Psychologie ist es keineswegs verwunderlich, dass die Psychologie bis heute weder über ein einheitliches Gegenstandsverständnis noch über ein einheitliches Wissensideal verfügt. Das wäre auch gar nicht wünschenswert, da ein solches Verständnis der Psychologie als Einheitswissenschaft die zulässigen Fragestellungen der Psychologie dogmatisch einschränken und den Reichtum an Perspektiven, unter denen die subjektive Welt des Menschen erforscht werden kann, zunichte machen würde.

Dass es in der Geschichte der Psychologie immer wieder bestimmte Schulrichtungen gegeben hat, die gleichsam einen Alleinvertretungsanspruch angemeldet haben, sei hier nur am Rande erwähnt. So wollte z.B. der Behaviorismus die Psychologie auf die "Wissenschaft vom Verhalten" (z.B. Morgan, 1961, 2) reduzieren, und Wilhelm Diltheys (1813–1911) Forderung, den Namen Psychologie für die "Wissenschaft vom sinnerfüllten Leben" (zit. n. Selg, 1975, 16) zu reservieren, findet bis heute Anhänger. In diesem Zusammenhang ist auch die immer wieder aufkommende Diskussion um das Menschenbild der Psychologie (z.B. Groeben & Erb, 2000) zu sehen, welche die naturwissenschaftlich geprägte Betrachtungsweise des Menschen als Organismus (oft fälschlich als mechanistisches Menschenbild bezeichnet) und die sozial- bzw. kulturwissenschaftliche Betrachtungsweise des Menschen als reflexives Subjekt einander gegenüberstellt.

Das Menschenbild
der Psychologie

Dabei sind zwei Diskursstränge miteinander vermischt: einerseits die Wiederbelebung des metaphysischen Dualismus von *Geist* und *Natur*,¹ und andererseits die Frage nach dem Erkenntnisinteresse (Habermas, 1968) der Psychologie bzw. die Frage danach, welches die Aufgaben psychologischer Wissensbildung sein sollen: technische Verwertbarkeit oder emanzipatorisches Reflexionswissen.

Erkenntnisinteresse
und Verwertungs-
zusammenhang

Eine Auseinandersetzung mit Fragen des Verwertungs zusammenhanges von Wissenschaft erweist sich aber nicht zuletzt deshalb als schwierig, weil das Erkenntnisinteresse der Psychologie zumeist nur von Kritikern der vorherrschenden Position in den Diskurs eingebracht wird, während die Vertreter des Mainstreams sich zumeist auf das Postulat der Wertfreiheit der Wissenschaften

1. Eine Vertiefung dieses Problems würde den Rahmen der vorliegenden Einführung sprengen. Siehe hierzu jedoch aus wissenschaftstheoretischer und kognitionswissenschaftlicher Sicht Carrier & Mittelstraß (1989).

zurückziehen und nur die Erkenntnis der Wirklichkeit als einzig legitimes Ziel der Wissenschaften gelten lassen. Insbesondere in der Grundlagenforschung ist diese Auffassung, wonach die Anwendung ihrer Forschungsergebnisse der Wissenschaft nur äußerlich ist, weit verbreitet.

Dabei wird jedoch übersehen, dass Gegenstandsverständnis und Wissensideal eines Forschungsprogramms nur relativ zu den jeweiligen Aufgaben der Wissensbildung (Lorenzen & Schwemmer, 1975) begründet werden können.

Aufgabenverständnis,
Gegenstandsverständnis
und Wissensideal

Unter dem Gegenstandsverständnis einer Wissenschaft ist die Antwort darauf zu verstehen, welche Art von Fragen in einer Wissenschaft als angemessen gelten. Mit dem Wissensideal meinen wir, welche Art von Antworten als zufriedenstellend erachtet werden (Hoyningen-Huene, 1989b), und unter dem Aufgabenverständnis verstehen wir die Antwort darauf, welche Art von Praxis das Wissen ermöglichen soll.

Sowohl ‚angemessen‘ als auch ‚zufriedenstellend‘ sind jedoch zweistellige Prädikatoren: Angemessen ist eine Frage immer nur relativ zu etwas, das selbst außerhalb dieser Frage liegt, und zufriedenstellend ist eine Antwort immer nur relativ zu etwas, das außerhalb der Antwort liegt. Dabei besteht zwischen Gegenstandsverständnis und Wissensideal insofern eine wechselseitige Abhängigkeit, als sich die Angemessenheit einer wissenschaftlichen Frage daran messen lässt, ob sie auf jene Art von Antworten abzielt, die dem jeweiligen Wissensideal entspricht; und ob eine Antwort zufriedenstellend ist, lässt sich daran messen, ob sie die jeweils gestellte Frage zu beantworten vermag. Wollten wir es dabei belassen, so würden wir uns entweder in einem definitiven Zirkel bewegen, oder wir müssten uns bezüglich des Wissensideals mit dem faktischen Konsens der Fachwissenschaftler zufrieden geben. Letzteres scheidet für die Psychologie aber schon deshalb aus, weil ein solcher Konsens hier de facto nicht existiert.

Die Orientierung des
Wissensideals an den
Aufgaben der
Wissensbildung

Ein Ausweg aus diesem Dilemma ergibt sich erst dann, wenn wir das Wissensideal an den jeweiligen Aufgaben der Wissensbildung ausrichten. So begründet z.B. Hoyningen-Huene (1989b) die Überlegenheit des Wissensideals der neuzeitlichen Naturwissenschaften damit, dass ihre technische Verwertbarkeit in diesem Wissensideal bereits strukturell angelegt ist. Eine szientistische Festlegung jeglicher Wissenschaft auf dieses Wissensideal lässt

sich daraus aber erst dann ableiten, wenn man die technische Verwertbarkeit des Wissens als einzig legitime Aufgabe der Wissensbildung festschreibt. Gerade dies aber ist es, was in der Psychologie zur Disposition steht.

Verzichtet man auf eine solche dogmatische Vorentscheidung, dann zeigt sich, dass in der Psychologie nicht nur de facto verschiedene Wissensideale nebeneinander bestehen, sondern auch, dass diese Wissensideale einander fruchtbar ergänzen. Wie sich bei genauerer Untersuchung ihrer (sach)logischen Struktur erweisen wird, implizieren diese verschiedenen Wissensideale jedoch unterschiedliche methodische Schritte zur Geltungsprüfung von Theorien.

Die wechselseitige Ergänzung verschiedener Wissensideale

Dabei kann in einem ersten Schritt zwischen einem aristotelischen und einem galileischen Wissensideal unterschieden werden.

Das aristotelische Wissensideal untersucht die wesentlichen Bestimmungsmerkmale einer bestimmten Klasse von Gegenständen und leistet insofern noch keine Erklärung der fraglichen Phänomene. Gleichwohl bildet es das unverzichtbare Fundament jeglicher Erfahrungswissenschaft, ohne das sich noch nicht einmal die Aufgaben der Wissensbildung bestimmen, geschweige denn begründen lassen.

Das aristotelische Wissensideal

Wenn wir z.B. die Aufgaben der (psychologischen) Friedensforschung in der Reduzierung von Gewalt mit gewaltfreien Mitteln (Galtung, 1993) sehen wollen, so müssen wir zu allererst geklärt haben, was wir mit Gewalt (engl. *violence*, von lat. *violare* = verletzen) meinen. U.a. müssen wir wissen, worin sich die Gewalt, die z.B. Fuchs (1993) als Verletzung mit mehr oder minder gravierenden Folgen definiert, etwa von der Aggression unterscheidet, die wir oben als (versuchte) Durchsetzung eigener Handlungen oder Ziele gegen den (tatsächlichen oder vermeintlichen) Willen eines anderen bestimmt haben.

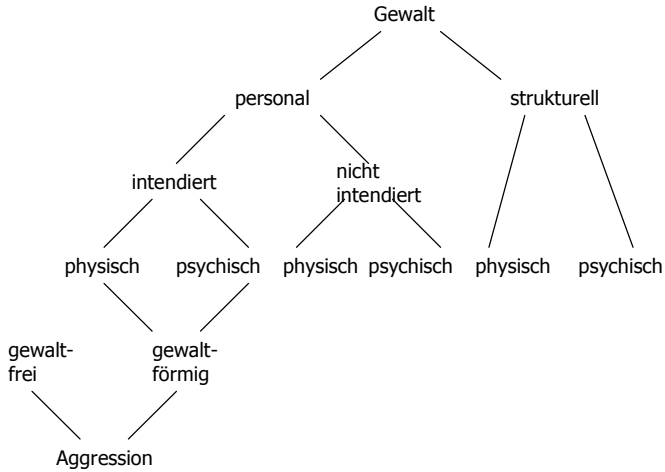
Beispiel:
Aufgaben der Friedensforschung

Obwohl wir damit zwar noch nicht erklärt haben, wie Gewalt entsteht, und obwohl wir daher auch noch kein Wissen darüber haben, wie Gewalt reduziert werden kann, haben wir mit der Festbeschreibung der wesentlichen Bestimmungsmerkmale von Aggression und Gewalt bereits ein Wissen über grundlegende Strukturmerkmale unseres Gegenstandsbereiches gebildet: namentlich, dass Aggression und Gewalt nicht einfach gleichgesetzt werden dürfen. Aggression ist zunächst nur zielstrebiges und tatkräf-

Wesentliche Bestimmungsmerkmale von Aggression und Gewalt

tiges Handeln – auch gegen den Willen eines anderen –, das noch lange nicht dessen Verletzung bedeuten muss. Gewalt dagegen besteht in eben dieser Verletzung und muss keineswegs das Ergebnis einer Aggression sein. Sie kann auch unbeabsichtigt oder strukturell bedingt sein.¹ Aggression und Gewalt überschneiden einander nur dort, wo sich die Durchsetzung eigener Ziele gewaltförmiger Mittel bedient (vgl. Abb. 1.4.1).

Abb. 1.4.1: Typologie des Aggressions- und Gewaltbegriffs (nach Kempf, 2000, 45)



Präzisierung der Erklärungsaufgabe

Damit ergibt sich zugleich eine Präzisierung unserer Erklärungsaufgabe: Reduzierung von Gewalt mit gewaltfreien Mitteln bedeutet nicht einfach die Unterdrückung von Aggression, sondern das Aggressionsproblem lautet nun, wie die gewaltförmige Eskalation von Konflikten vermieden bzw. rückgängig gemacht werden kann.

Das galileische Wissensideal

Im Unterschied zum aristotelischen bietet das galileische Wissensideal eine *Erklärung* der fraglichen Phänomene an, indem es Relationen zwischen verschiedenen Klassen von Gegenständen (z.B. Ursache und Wirkung oder Ziel und Mittel) untersucht. Dabei lassen sich innerhalb des galileischen Wissensideals nochmals verschiedene Erklärungsmodelle unterscheiden (vgl. Abb. 1.4.2).

1. Zur Unterscheidung zwischen personaler und struktureller Gewalt vgl. Galtung (1969).

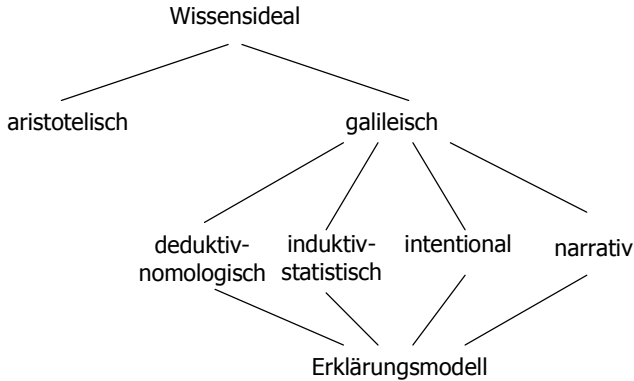


Abb. 1.4.2: Überblick über verschiedene Wissensideale und Erklärungsmodelle, welche in der Psychologie vertreten sind

Als Vorbild diente der Psychologie zunächst das deduktiv-nomologische Erklärungsmodell der Newton'schen Mechanik, welche die Ursache-Wirkungs-Relation in Form von Naturgesetzen darstellt. Als Beispiel hierfür kann das Weber'sche Gesetz dienen: Die Ursache dafür, dass z.B. beim Auflegen auf den Handrücken der Unterschied zweier Gewichte von 100g und 110g nicht wahrgenommen wird, wird durch das Weber'sche Gesetz dahingehend erklärt, dass dieser Gewichtsunterschied unterhalb der relativen Unterschiedsschwelle von $1/3$ des Standardreizes liegt.

Die deduktiv-nomologische Erklärung

Nicht nur in der Wahrnehmungspsychologie, sondern auch in anderen Teilgebieten der Psychologie hat man immer wieder versucht, solche allgemeingültigen Gesetzmäßigkeiten zu finden. So postulierten z.B. Dollard, Doob, Miller, Mowrer & Sears (1939) in der ersten Fassung der sogenannten Frustrations-Aggressions-Hypothese, dass Aggression stets eine Folge von Frustration sei sowie, dass Frustration stets zu einer Form von Aggression führe.¹

In dieser streng deterministischen Form ließ sich die Frustrations-Aggressions-Hypothese jedoch nicht aufrechterhalten und wurde schon bald durch Miller (1941) und Sears (1941) dahingehend abgeschwächt, dass Frustration Anreize zu verschiedenen Arten

Die induktiv-statistische Erklärung

1. Um Begriffsverwirrung zu vermeiden, ist allerdings darauf hinzuweisen, dass Dollard et al. unter Aggression eine Verhaltenssequenz verstehen, die auf die Verletzung eines Organismus oder Organismusersatzes abzielt, und das so definierte aggressive Verhalten dann – konsequent behavioristisch – durch seine (gewaltförmigen) Folgen operationalisieren. Aus Gründen der terminologischen Konsistenz sprechen wir daher im Kontext der Frustrations-Aggressions-Hypothese statt von Aggression besser von gewaltförmigem Verhalten.

von Verhaltensweisen bietet, unter denen sich stets auch ein Anreiz zu gewaltförmigem Verhalten befindet. Die von Dollard et al. postulierte Gesetzmäßigkeit wurde damit probabilistisch abgeschwächt und die deduktiv-nomologische Erklärung durch eine induktiv-statistische Erklärung ersetzt. Gewaltförmiges Verhalten ist demnach nicht mehr eine unausweichliche Wirkung der Frustration, sondern Frustration ist lediglich eine Bedingung, unter welcher gewaltförmiges Verhalten wahrscheinlich(er) wird.

Wie man heute weiß, gilt auch das Weber'sche Gesetz nur im Bereich mittlerer Reizintensitäten, und selbst dort nur ungefähr. Empirische Gesetzmäßigkeiten gelten in der Psychologie einfach nicht mit derselben Strenge wie in den ‚harten‘ Naturwissenschaften, und das deduktiv-nomologische Erklärungsmodell wurde in der Psychologie inzwischen generell durch das induktiv-statistische Erklärungsmodell verdrängt.

Mit dieser Abschwächung des Wissensideals ging zugleich auch eine Abschwächung des Gegenstandsverständnisses der Psychologie einher. Denn während die deduktiv-nomologische Erklärung auf die Frage nach den Ursachen des zu erklärenden Phänomens antwortet, antwortet die induktiv-statistische Erklärung nur noch auf die Frage nach den Bedingungen, unter welchen mit ihm (mehr oder minder wahrscheinlich) zu rechnen ist.

Die intentionale Erklärung

Eine radikale Erweiterung des Gegenstandsverständnisses der Psychologie impliziert die Frage nach den Gründen, warum sich jemand so und nicht anders verhält. Denn die Frage nach den je subjektiven Handlungsgründen zielt auf eine intentionale Erklärung, welche nach dem Schema der Ziel-Mittel-Relation zu konstruieren ist und welche damit einer gänzlich anderen Logik folgt: Im Unterschied zur deduktiv-nomologischen Erklärung und ihrer induktiv-statistischen Abschwächung bewegt sich die intentionale Erklärung nicht mehr in einem Bedingungsdiskurs, sondern in einem Begründungsdiskurs.

Im Schulenstreit zwischen naturwissenschaftlich und sozialwissenschaftlich orientierten Psychologen werden diese beiden Diskursformen mitunter als unversöhnliche Gegensätze einander gegenübergestellt (z.B. Holzkamp, 1986). Tatsächlich sind sie jedoch oft miteinander verschränkt, wie sich ebenfalls am Beispiel der Frustrations-Aggressions-Hypothese veranschaulichen lässt, denn Frustration wird von Dollard et al. als Unterbrechung einer zielgerichteten Aktivität definiert.

Auch wenn wir z.B. die Frage stellen, warum ein gegebener Konflikt plötzlich gewaltförmig eskaliert ist, wird sich darauf keine zufriedenstellende Antwort finden lassen, die sich ausschließlich einer der beiden Diskursformen bedient. Worauf Fragen dieser Art abzielen, ist die Herstellung einer sinnstiftenden Verbindung zwischen zwei perspektivischen Eckpunkten (vorher vs. nachher), wie sie im sog. narrativen Erklärungsmodell dargestellt wird.

Die narrative Erklärung

Dass es im Grenzbereich zwischen Psychologie und Biologie eine Vielzahl an Fragestellungen gibt, die sich mit rein naturwissenschaftlichen Methoden bearbeiten lassen, bleibt davon ebenso unberührt wie deren Legitimität. Wollte man Psychologie jedoch auf den Bereich dieser Fragestellungen einschränken, so ließe sich nicht mehr verständlich machen, warum es neben der Biologie auch noch die Psychologie als eigene Wissenschaft geben soll.

Indem die Psychologie (notgedrungen) darüber hinaus geht, muss sie sich allerdings der Frage stellen, ob sie ihren Anspruch auf Erfahrungswissenschaftlichkeit auch tatsächlich einzulösen vermag. Gerade dies ist es, was psychologische Forschungsmethoden so spannend macht, denn, wie Heinz von Foerster (1972) einmal festgestellt hat:

"Die ‚hard sciences‘ sind erfolgreich, weil sie sich mit den ‚soft problems‘ beschäftigen; die ‚soft sciences‘ haben zu kämpfen, denn sie haben es mit den ‚hard problems‘ zu tun".

Für die Psychologie als Substanzwissenschaft mag es schmerzhaft sein, dass sich psychische Krankheiten nicht so einfach "reparieren" lassen wie ein defektes Automobil, oder dass die Lösung von Konflikten nicht so einfach zu bewerkstelligen ist wie ein Flug zum Mond. Für die Methodenlehre als Teildisziplin der Psychologie sind es aber gerade die ‚hard problems‘ einer ‚soft science‘, die ihren Sex-Appeal ausmachen.

Kontrollfragen

- Was versteht man unter
 - a) dem Aufgabenverständnis,
 - b) dem Gegenstandsverständnis und
 - c) dem Wissensideal einer Wissenschaft?
- Welche Abhängigkeit besteht zwischen Gegenstandsverständnis und Wissensideal?
- Wie kann man vermeiden, dass man sich bei der Festlegung von Ge-

gegenstandsverständnis und Wissensideal in einen definitorischen Zirkel verstrickt?

- Worin besteht der Unterschied zwischen dem aristotelischen und dem galileischen Wissensideal?
- Worin besteht der Unterschied zwischen Aggression und Gewalt?
- Welche grundlegenden Strukturmerkmale des Gegenstandsbereiches der (psychologischen) Friedensforschung ergeben sich aus dieser Unterscheidung?
- Welche Präzisierung der Erklärungsaufgabe der Friedensforschung ergibt sich daraus?
- Nennen Sie vier Erklärungsmodelle, welche sich innerhalb des galileischen Wissensideals unterscheiden lassen!
- Durch welches dieser Erklärungsmodelle wird
 - a) die Ursache-Wirkungs Relation,
 - b) die Ziel-Mittel Relationdargestellt?
- Erläutern Sie am Beispiel der Frustrations-Aggressions-Hypothese,
 - a) inwiefern die statistisch-induktive Erklärung eine Abschwächung der deduktiv-nomologischen Erklärung darstellt sowie
 - b) welche Abschwächung des Gegenstandsverständnisses der Psychologie damit einhergeht!

Kapitel 2

Hard Science: Das Erbe der Naturwissenschaften

2.1

Das aristotelische Wissensideal

Das aristotelische Wissensideal ist das Ideal der beweisenden Wissenschaft nach dem Vorbild der euklidischen Geometrie. Am Anfang stehen erste Sätze, die sogenannten Axiome, die selbst nicht beweisbar – wohl aber begründbar – sind und die die wesentlichen Eigenschaften einer bestimmten Klasse von Gegenständen beschreiben. In der euklidischen Geometrie sind dies z.B. die Punkte, Geraden und Ebenen des dreidimensionalen euklidischen Raumes, deren wesentliche Eigenschaften u.a. durch die folgenden Axiome beschrieben werden (vgl. Meschkowski, 1966):

Das Vorbild der euklidischen Geometrie

A1: Zu jeder Geraden gehören mindestens zwei Punkte, die auf der Geraden liegen.

A2: Zu irgendzwei verschiedenen Punkten gehört genau eine Gerade, auf der die beiden Punkte liegen.

A3: Zu jeder Ebene gehören mindestens drei verschiedene, nicht in einer Geraden liegende Punkte, die in der Ebene liegen.

Aus diesen ersten Sätzen werden weitere Sätze logisch deduziert (= abgeleitet). Z.B. folgt unmittelbar aus Axiom A2 der folgende:

Satz: Zwei verschiedene Geraden haben einen oder keinen Punkt gemeinsam.

Beweis: Nach Axiom A2 können zwei verschiedene Geraden nicht zwei verschiedene Punkte gemeinsam haben.

Dass sich zwei verschiedene Geraden im dreidimensionalen euklidischen Raum in genau einem Punkt oder überhaupt nicht schneiden, ist somit keine empirische, sondern eine strukturelle Gesetzmäßigkeit, die sachlogisch begründet ist. Wir können daher auch sagen: Das aristotelische Wissensideal zielt auf die sachlogische Begründung struktureller Gesetzmäßigkeiten des jeweiligen Gegenstandsbereiches ab.

Sachlogische Begründung struktureller Gesetzmäßigkeiten

Dieses ist das Wissensideal der sogenannten Formalwissenschaften, insbesondere der Mathematik, hat dort aber gewisse Abstriche erfahren, auf die wir später zu sprechen kommen. Indem sie sich der Mathematik als Instrument bedienen, bestehen aber auch die Theorien der Naturwissenschaften nicht nur aus einer Aneinanderreihung von Naturgesetzen, sondern sie enthalten strukturelle Theorieanteile.

Auch in der Psychologie gibt es eine Reihe von Theorien, die strukturelle Gesetzmäßigkeiten untersuchen, z.B. die allgemeine Messtheorie (Suppes & Zinnes, 1963; Pfanzagl, 1968; einführend: Orth, 1974), die von Smedslund (1988) begründete Psycho-Logik oder die erstmals von Gulliksen (1950) axiomatisch ausgearbeitete klassische Testtheorie (vgl. dazu einführend: Lord & Novick, 1968; Fischer, 1974; Steyer & Eid, 1993).

2.1.1 Die axiomatische Begründung der klassischen Testtheorie

Zuverlässigkeit
psychologischer
Testergebnisse

Seit der Konstruktion des ersten Intelligenztests durch Binet & Simon (1905) stellt die Frage nach der Verlässlichkeit psychologischer Testergebnisse ein Problem von nicht unerheblicher praktischer Bedeutung dar. Zum einen könnte die Testleistung eines Probanden durch Störfaktoren (z.B. Müdigkeit) beeinträchtigt werden; zum anderen wäre es denkbar, dass Tests bestimmte Probanden (z.B. Hochbegabte) systematisch bevorzugen und andere (z.B. Minderbegabte) systematisch benachteiligen usw. Kurz und gut: Es sind Zweifel angebracht, wie zuverlässig sich die Leistungsfähigkeit der Probanden in den Testergebnissen widerspiegelt.

Um diese Frage fundiert beantworten zu können, hat man schon früh damit begonnen, statistische Rechenformeln zu entwickeln, welche es erlauben sollten, die Qualität von Tests bzw. Testergebnissen empirisch zu untersuchen. Dazu sind verschiedene Verfahren vorgeschlagen worden, z.B. die Retest-Reliabilität. Zu ihrer Bestimmung werden die Probanden zweimal mit demselben Test untersucht. Die Übereinstimmung zwischen den Ergebnissen der beiden Testuntersuchungen wird dann als Maß für die Messgenauigkeit des Tests verwendet. Ein anderes Verfahren ist die Paralleltest-Reliabilität, bei der die Probanden stattdessen mit zwei gleichwertigen Tests untersucht werden usw.

Gulliksen (1950) hat als erster den Versuch unternommen, diese Rechenformeln in eine geschlossenen Theorie psychologischer Testergebnisse zu integrieren. Ausgangspunkt dieser sog. klassischen Testtheorie ist die Annahme, dass das Testergebnis, der Score (x_{vt}), den ein Proband (v) in einem Test (t) erzielt, fehlerbehaftet sein kann. D.h. das empirisch vorliegende Testergebnis des Probanden muss nicht mit dem Wert übereinstimmen, der ihm eigentlich zusteht. Letzterer wird als wahrer Testwert oder True-Score (τ_{vt}) bezeichnet. Der Messfehler des Testergebnisses (f_{vt}) wird dann als die Abweichung des Testergebnisses vom True-Score des Probanden definiert ($f_{vt} = x_{vt} - \tau_{vt}$). Daraus ergibt sich die Grundgleichung der klassischen Testtheorie, wonach das Testergebnis aus True-Score und Messfehler zusammengesetzt ist:

$$x_{vt} = \tau_{vt} + f_{vt}$$

Darauf aufbauend, formuliert Gulliksen vier Axiome:

1. Der mittlere Messfehler ist gleich Null.
2. Der Messfehler korreliert nicht mit dem True-Score.¹
3. Die Messfehler aus verschiedenen Tests korrelieren nicht miteinander.
4. Der Messfehler in einem Test korreliert nicht mit dem True-Score in einem anderen Test.

Wie Gulliksen zeigt, können aus diesen Axiomen tatsächlich die gesamten Ergebnisse der klassischen Testtheorie hergeleitet werden.

So kann u.a. auch gezeigt werden, dass Retest-Reliabilität und Paralleltest-Reliabilität dasselbe Maß für die Messgenauigkeit des Tests berechnen²: die Reliabilität des Tests (ρ^2_{XT}), die als systematischer Varianzanteil der Testergebnisse definiert ist.³

Die Reliabilität ist demnach eine Zahl zwischen Null (0%) und Eins (100%). Beträgt die Reliabilität Null, so sind die gesamten Unterschiede zwischen den Testergebnissen auf Messfehler zurückzuführen und spiegeln keinerlei (0%) tatsächliche Leistungs-

Die Grundgleichung der klassischen Testtheorie

Die Axiome von Gulliksen

Die Reliabilität

1. Die Korrelation ist ein statistisches Maß, welches die wechselseitige Abhängigkeit zweier Messreihen beschreibt.
2. Unter bestimmten Voraussetzungen bezüglich der Gleichwertigkeit von Test und Testwiederholung bzw. von Test und Paralleltest.
3. Die Varianz ist ein statistisches Maß, welches die Streuung einer Variable um ihren Mittelwert beschreibt.

unterschiede zwischen den Probanden wider. Beträgt die Reliabilität Eins, so geht die Streuung der Testergebnisse dagegen zu 100% auf tatsächliche Leistungsunterschiede zurück, und der Messfehler hat keinerlei Anteil daran.

2.1.2 Anforderungen an Axiomensysteme

Widerspruchsfreiheit,
Unabhängigkeit und
Vollständigkeit

Mit seinem 1899 erschienenen Buch *Grundlagen der Geometrie* schuf David Hilbert die Voraussetzung für metamathematische Untersuchungen über die Widerspruchsfreiheit und Unabhängigkeit von Axiomensystemen. Die von Hilbert (zit.n. Meschkowski, 1966, 9) gestellten Anforderungen an ein brauchbares Axiomensystem lauten:¹

1. Das System soll widerspruchsfrei sein.
2. Die Sätze des Axiomensystems sollen voneinander unabhängig sein.
3. Das Axiomensystem soll vollständig sein.

Verzicht auf die
Definition der
Grundbegriffe

Im Unterschied zu Euklid (ca. 365–ca. 300 v. Chr.), der seine "Elemente" auf Definitionen der Grundbegriffe sowie auf fünf Postulate und neun Axiome gegründet hatte,² verzichtete Hilbert jedoch auf eine Definition der Grundbegriffe der Geometrie (Punkt, Gerade, Ebene) ebenso wie auf jegliche Anschauung. Am Anfang seiner *Grundlagen der Geometrie* stehen lediglich die Sätze:

"Wir denken uns drei verschiedene Systeme von Dingen: Die Dinge des ersten Systems nennen wir Punkte und bezeichnen sie mit A, B, C, \dots ; die Dinge des zweiten Systems nennen wir Geraden und bezeichnen sie mit a, b, c, \dots ; die Dinge des dritten Systems nennen wir Ebenen und bezeichnen sie mit α, β, γ ; die Punkte heißen auch Elemente der linearen Geometrie, die Punkte und Geraden heißen die Elemente der ebenen Geometrie, und die Punkte, Geraden und Ebenen heißen die Elemente der räumlichen Geometrie oder des Raumes. Wir denken die Punkte, Geraden und Ebenen in gewissen gegenseitigen Beziehungen und bezeichnen diese Beziehungen durch Worte wie ‚liegen‘, ‚zwischen‘, ‚parallel‘, ‚kongruent‘, ‚stetig‘; die genaue und vollständige Beschreibung dieser Be-

1. Während die erste dieser Forderungen unbestritten blieb, hat man die Forderung nach Vollständigkeit von Axiomensystemen (z.B. in der mathematischen Gruppen- und Verbandstheorie) teilweise fallengelassen. An die Stelle der zweiten Forderung tritt in der modernen Axiomatik häufig die Forderung, dass ein Axiomensystem möglichst wenige und möglichst einfache Axiome enthalten soll (vgl. Meschkowski, 1966, 16).
2. Die Unterscheidung zwischen Postulaten und Axiomen ist heute nicht mehr üblich.

ziehungen erfolgt durch die „Axiome der Geometrie“ (Hilbert, 1898, zit.n. Meschkowski, 1986, 54).

Mit dieser Auffassung, wonach die Grundbegriffe der Mathematik nicht zu definieren sind, legte Hilbert den Grundstein für das moderne Selbstverständnis der Mathematik, das die Mathematik als die Wissenschaft von den formalen Systemen (oder Kalkülen) ansieht. Sie selbst sagt über die uns umgebende Wirklichkeit nichts aus. Welche formalen Systeme von den Erfahrungswissenschaften (z.B. der Physik) benutzt werden, ist eine Frage der Zweckmäßigkeit.

Mathematik als
Formalwissenschaft

Hilbert selbst (zit.n. Meschkowski, 1986, 55) hat diesen Standpunkt einmal sehr drastisch formuliert, indem er gesagt hat: "Man muss jederzeit an Stelle von ‚Punkten, Geraden und Ebenen‘, ‚Tische, Stühle, Bierseidel‘ sagen können".

Damit hat Hilbert zwei Dinge gleichzeitig angesprochen, die man besser auseinanderhalten sollte:

Die Irrelevanz der
Lautgestalt von
Termini

Erstens, dass es auf die Lautgestalt der Termini einer Wissenschaftssprache überhaupt nicht ankommt, und zwar weder in der Mathematik noch in den Erfahrungswissenschaften. Entsprechend könnten wir auch die oben verwendeten Wörter ‚Aggression‘ und ‚Gewalt‘ gegeneinander austauschen, sie durch die Wörter ‚Durchsetzungshandlung‘ und ‚Verletzung‘ ersetzen, oder genauso gut von ‚Surfbrettern‘ und ‚Skateboards‘ reden, so lange wir nur die mit dieser Terminologie getroffene sprachliche Unterscheidung konsequent einhalten und uns nicht etwa durch die Lautgestalt der Termini dazu verleiten lassen, mehr oder anderes in sie hineinzulesen, als terminologisch vereinbart wurde.

So dürfen wir z.B. nicht sagen, dass die Unterbrechung einer zielgerichteten Aktivität Frustration *bewirkt*. In der Terminologie der Frustrations-Aggressions-Hypothese *ist* die Unterbrechung der zielgerichteten Aktivität die Frustration (lat. *frustrare* = vereiteln), und die dadurch ausgelösten Gefühle (die sich irgendwo zwischen Enttäuschung, Wut und Angriffslust bewegen können) sind davon klar zu unterscheiden – auch wenn wir (alltagssprachlich) diesen Gefühlsmix gerne mit der neudeutschen Wendung umschreiben: "Ich bin frustriert".

Ein Beispiel aus der
Psychologie

Arbeitsteilung
zwischen Formal-
und Erfahrungswissenschaften

Zweitens hat Hilbert mit seiner drastischen Formulierung eine Position eingenommen, welche die Mathematik für die Begründung der Axiome und die Klärung der darin vorkommenden Grundbegriffe als unzuständig erklärt. Die Ausarbeitung dieses Aspekts des aristotelischen Wissensideals ist Angelegenheit der Erfahrungswissenschaften bzw. ihrer prototheoretischen Grundlegung.¹ Anders ausgedrückt: Die Mathematik kümmert sich nur um die strukturellen Gesetzmäßigkeiten, die innerhalb eines formalen Systems herrschen, und wenn die Erfahrungswissenschaften sich eines bestimmten mathematischen Kalküls bedienen, so haben sie selbst den Nachweis zu erbringen, dass der jeweilige Gegenstandsbereich ein Modell für diesen Kalkül ist.²

Prototheoretische
Begründungen

Letzteres impliziert aber z.B. zur Grundlegung der physikalischen Längenmessung die Notwendigkeit einer Protogeometrie, in der es um genau jene Grundlagenfragen geht, für die sich die moderne Mathematik nicht (mehr) als zuständig erklärt. Vereinfacht ausgedrückt, handelt die Protogeometrie davon, wie Gegenstände der wirklichen Welt (Punkte, Geraden, Ebenen) so konstruiert werden können, dass die Axiome der euklidischen Geometrie für sie zutreffen. Wie dies im einzelnen geschieht, kann hier nicht näher ausgeführt werden und würde wohl auch zu weit vom Gegenstand des vorliegenden Buches wegführen.³ Exemplarisch für prototheoretische Argumentationen werden wir jedoch weiter unten die konstruktive Definition des Wahrscheinlichkeitsbegriffs kennen lernen, der in der psychologischen Methodenlehre eine zentrale Rolle einnimmt.

Wie am Beispiel der klassischen Testtheorie veranschaulicht werden kann, handelt es sich bei diesen Grundlagenfragen nicht bloß um Probleme, die Philosophen oder Wissenschaftstheoretiker interessieren mögen, sondern es geht um ganz konkrete Methodenprobleme mit weitreichenden Konsequenzen für die Erfahrungswissenschaften selbst.

-
1. Als Prototheorie bezeichnet man im Rahmen der konstruktivistischen Wissenschaftstheorie die einer theoriefähigen Bemühung methodisch vorausgehende Vor-Theorie (vgl. Janich, 1995).
 2. Eine Menge von Dingen heißt ein Modell für ein Axiomensystem S , wenn für sie die Aussagen von S zutreffen, ihre Existenz aber bereits anderweitig gesichert ist (vgl. Meschkowski, 1966, 102).
 3. Siehe hierzu die sehr anschauliche Darstellung in dem Lehrbuch von Lorenzen (1987, 191ff).

2.1.3 Kritik und Neubegründung der klassischen Testtheorie

Indem Gulliksen auf eine Definition des True-Scores verzichtet und in seinen Axiomen lediglich bestimmte Eigenschaften der statistischen Verteilung von True-Scores und Messfehlern fest schreibt, orientiert er sich streng an Hilberts Vorgaben mit all ihren Konsequenzen. Für die mathematische Herleitung der Rechenformeln der klassischen Testtheorie ist das auch völlig ausreichend. Für den psychologischen Diagnostiker, der die Qualität von Testergebnissen beurteilen will, ist es aber – um in Hilberts Bild zu bleiben – keineswegs gleichgültig, ob Gulliksen Theorie von True-Scores und Messfehlern handelt oder von Schanktischen und Bierseideln. Anders ausgedrückt: Für den Diagnostiker ist es entscheidend, ob Gulliksen "Bierseidel" ($f_{\sqrt{t}}$) irgendetwas mit dem zu tun haben, was er im Blick hat, wenn er von der (möglichen) Fehlerbehaftetheit psychologischer Testergebnisse spricht. Für den Diagnostiker stellt sich die Frage, ob es überhaupt zweckmäßig ist, in der psychologischen Testtheorie von Gulliksen Axiomen auszugehen bzw. ob eine Testtheorie, die auf diesen Axiomen beruht, überhaupt geeignet sein kann, jene Fragen zu beantworten, die sich für den Diagnostiker bezüglich der Verlässlichkeit psychologischer Testergebnisse stellen.

Die Frage nach der Zweckmäßigkeit der Axiome von Gulliksen

Zweifel daran sind schon insofern angebracht, als Gulliksen Axiome ja bestimmte wünschenswerte Eigenschaften psychologischer Tests festschreiben. So garantiert z.B. das erste Axiom, dass die Testergebnisse keinen systematischen Messfehler enthalten; das zweite Axiom schließt aus, dass z.B. die Testleistung hochbegabter Probanden überschätzt wird, während minderbegabte durch den Messfehler noch zusätzlich benachteiligt werden usw.

Das Erfordernis der Modellkonstruktion

Für die Brauchbarkeit der Testtheorie als Instrument zur kritischen Beurteilung von Tests und Testergebnissen muss sich dies verheerend auswirken. Denn wenn diese Qualitätsmerkmale von Tests bereits a priori vorausgesetzt werden, dann ist eine Testtheorie, die auf diesen Voraussetzungen aufbaut, von vornherein ungeeignet dafür, mögliche Verletzungen dieser Qualitätsmerkmale aufzuzeigen. Damit ergibt sich der – z.B. von Grubitzsch & Rexilius (1978) formulierte – Verdacht, dass die klassische Testtheorie letztlich nur eine Immunisierung psychologischer Tests gegenüber Kritik leistet.

Dieser Verdacht kann erst dann ausgeräumt werden, wenn sich erstens ein sinnvolles psychologisches Modell der klassischen Testtheorie angeben lässt, und wenn zweitens die Termini ‚True-Score‘ und ‚Messfehler‘ nur in dem Sinne verwendet werden, in dem sie durch dieses Modell definiert sind.

Die Grundannahmen von Novick

Ein solches Modell wurde von Novick (1966) konstruiert. Die Modellvoraussetzungen lauten:

1. Jeder Testung (t) eines Probanden (v) entspricht eine zufällige Variable möglicher Testergebnisse (X_{vt}) mit endlichem Erwartungswert $E(X_{vt})$ und endlicher Varianz $\sigma^2(X_{vt})$. Diese nennen wir die Scorevariable.¹
2. Das Testergebnis (x_{vt}), welches der Proband erzielt hat, ist eine unabhängige Realisation dieser Scorevariable.
3. Der True-Score des Probanden (τ_{vt}) ist per definitionem gleich dem Erwartungswert der Scorevariable: $\tau_{vt} = E(X_{vt})$.

Geht man von diesen Modellvoraussetzungen aus, dann stehen die Axiome von Gulliksen bereits als bewiesene Sätze zur Verfügung (vgl. Kapitel 2.1.4). Gleichzeitig wird mit der expliziten Definition des True-Scores als Erwartungswert der Scorevariable aber auch deutlich, dass der Messfehler der klassischen Testtheorie ausschließlich reine Zufallsfehler beschreibt, welche sich daraus ergeben, dass einem Probanden eben nicht ein bestimmtes Testergebnis entspricht, sondern eine Wahrscheinlichkeitsverteilung möglicher Testergebnisse, die um den – als True-Score bezeichneten – Erwartungswert dieser Verteilung streuen. Darüber hinausgehende Fehlerquellen (z.B. Störfaktoren wie Müdigkeit oder die Bevorzugung hochbegabter Probanden durch den Test) sind schlichtweg nicht Gegenstand der klassischen Testtheorie, und die Analyse von Tests mittels der klassischen Testtheorie sagt über sie nichts aus.

1. Als Erwartungswert bezeichnet man den Mittelwert einer Wahrscheinlichkeitsverteilung.

2.1.4 Exkurs: Herleitung der Axiome von Gulliksen¹

Gegeben sei eine Zufallsvariable T_{ot} , die auf einer Personenpopulation P definiert ist und die Werte τ_{vt} (True-Score) annimmt. Jeder V_p ($v = 1, 2, 3, \dots$) aus P entspreche zudem eine Zufallsvariable X_{vt} mit endlichem Erwartungswert $E(X_{vt}) = \tau_{vt}$ und endlicher Varianz $\sigma^2(X_{vt})$. Jede Testung einer V_p v sei eine unabhängige Realisation x_{vt} dieser Zufallsvariablen. Dann stellt die Differenz $f_{vt} = x_{vt} - \tau_{vt}$ (Messfehler) die Realisation einer Zufallsvariable $F_{vt} = X_{vt} - \tau_{vt}$ dar, mit dem Erwartungswert

$$E(F_{vt}) = 0 \quad [1]$$

und der Varianz

$$\sigma^2(F_{vt}) = \sigma^2(X_{vt}) \quad [2]$$

Voraussetzung: die Grundannahmen von Novick

Beweis der Gleichungen [1] und [2]:

Ad Gleichung [1]: Wegen $F_{vt} = X_{vt} - \tau_{vt}$ und $E(X_{vt}) = \tau_{vt}$ gilt für den Erwartungswert des Messfehlers

Der Erwartungswert des Messfehlers

$$E(F_{vt}) = E(X_{vt} - \tau_{vt}) = \tau_{vt} - \tau_{vt} = 0.$$

Ad Gleichung [2]: Da der True-Score für eine feste V_p v eine Konstante ist, gilt für die Varianz des Messfehlers

Die Varianz des Messfehlers

$$\sigma^2(F_{vt}) = \sigma^2(X_{vt} - \tau_{vt}) = \sigma^2(X_{vt}).$$

Seien t und t' zwei Tests, die den genannten Voraussetzungen genügen. Dann gelten für die Verteilung der Messfehler über die Personenpopulation hinweg die Axiome von Gulliksen:

Behauptung: die Axiome von Gulliksen

1. $E(F_{ot}) = 0$
2. $\rho(F_{ot}, T_{ot}) = 0$
3. $\rho(F_{ot}, F_{ot'}) = 0$
4. $\rho(F_{ot}, T_{ot'}) = 0$

Herleitung der Axiome 1 und 2:

Beweise

Axiom 1: Der mittlere Messfehler ist gleich Null. Formal:

Axiom 1

$$E(F_{ot}) = 0.$$

1. Dieser Abschnitt kann ohne Verlust der Kontinuität übersprungen werden.

Beweis: Für den mittleren Messfehler in der Personenpopulation gilt

$$E(F_{ot}) = E_V\{E(F_{vt})\},$$

worin $E_V\{\dots\}$ eine Erwartungswertbildung über die Vpn hinweg bezeichnet. Wegen Gleichung [1] sind die $E(F_{vt})$ aber alle gleich Null, und, da der Erwartungswert einer Konstanten gleich dieser Konstanten ist, folgt:

$$E(F_{ot}) = 0.$$

Axiom 2

Axiom 2: Der Messfehler korreliert nicht mit dem True-Score. Formal:

$$\rho(F_{otr}, T_{ot}) = 0.$$

Beweis: Da der Erwartungswert $E(F_{vt})$ laut Gleichung [1] für jede beliebige Vp v gleich Null ist, gilt insbesondere für alle jene Vpn, die den True-Score τ_{vt} haben, die Gleichung

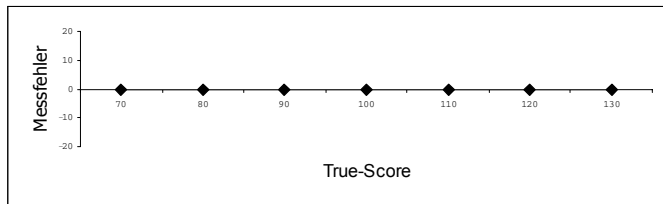
$$E(F_{vt}|\tau_{vt}) = 0$$

und daher für alle möglichen True-Scores:

$$E(F_{ot}|T_{ot} = \tau_{vt}) = 0.$$

Die wahre Regression des Messfehlers auf den True-Score fällt daher mit der Abszisse zusammen (vgl. Abb. 2.1.1): Sie ist linear mit dem Anstieg Null. Folglich verschwindet die Korrelation zwischen Messfehler und True-Score.

Abb. 2.1.1: Regression des Messfehlers auf den True-Score



Die Axiome 3 und 4 werden analog hergeleitet:

Axiom 3

Axiom 3: Die Messfehler aus verschiedenen Tests t und t' korrelieren nicht miteinander. Formal:

$$\rho(F_{otr}, F_{ot'}) = 0.$$

Beweis: Die Korrelation zwischen den Messfehlern aus zwei verschiedenen Tests t und t' verschwindet, weil Gleichung [1] insbesondere für all jene V_{pn} gilt, deren Testergebnis in Test t' mit dem Messfehler $f_{vt'}$ belastet ist, so dass $E(F_{vt}|f_{vt'}) = 0$ und daher für alle möglichen Messfehler in Test t' : $E(F_{ot}|F_{ot'} = f_{vt'}) = 0$.

Axiom 4: Der Messfehler in einem Test t korreliert nicht mit dem True-Score in einem anderen Test t' . Formal: Axiom 4

$$\rho(F_{ot}, T_{ot'}) = 0.$$

Beweis: Die Korrelation zwischen dem Messfehler in Test t und dem True-Score in Test t' verschwindet, weil Gleichung [1] insbesondere für all jene V_{pn} gilt, die in Test t' den True-Score $\tau_{vt'}$ haben, so dass $E(F_{vt}|\tau_{vt'}) = 0$ und daher für alle möglichen True-Scores in Test t' : $E(F_{ot}|T_{ot'} = \tau_{vt'}) = 0$.

Spätestens aus der Herleitung des Axioms 4 wird erkennbar, dass der Messfehler der klassischen Testtheorie ausschließlich Zufallsfehler beschreibt, nicht jedoch Störeinflüsse wie Müdigkeit oder mangelnde Aufmerksamkeit, die vielmehr dem True-Score zugeschlagen werden.

Zur Interpretation
des Messfehlers

Zum Beweis genügt ein einfaches Gegenbeispiel: Angenommen, Test t ist ein Intelligenztest. Test t' misst die Konzentrationsleistung. Je schlechter die Aufmerksamkeit einer V_p , desto schlechter schneidet sie im Intelligenztest ab, so dass die Regression der Intelligenztestleistung (X_{ot}) auf die Aufmerksamkeit ($T_{ot'}$) einer monoton wachsenden Funktion

$$E(X_{ot}|T_{ot'} = \tau_{vt'}) = f(\tau_{vt'})$$

folgt. Wegen $X_{ot} = T_{ot} + F_{ot}$ ist aber

$$E(X_{ot}|T_{ot'} = \tau_{vt'}) = E(T_{ot}|T_{ot'} = \tau_{vt'}) + E(F_{ot}|T_{ot'} = \tau_{vt'}),$$

und da gemäß der Herleitung von Axiom 4 der Erwartungswert $E(F_{ot}|T_{ot'} = \tau_{vt'}) = 0$, folgt

$$E(T_{ot}|T_{ot'} = \tau_{vt'}) = f(\tau_{vt'}).$$

Also schlägt sich die Störvariable „Aufmerksamkeit“ nicht auf den Messfehler, sondern auf die tatsächliche Leistungsfähigkeit (True-Score) der Probanden im Intelligenztest nieder.

Entsprechend beschreibt der True-Score der klassischen Testtheorie auch nicht *die* Leistungsfähigkeit eines Probanden in dem jeweiligen Test schlechthin, sondern lediglich seine Leistungsfähig-

Zur Interpretation
des True-Scores

keit unter den jeweiligen Versuchs- und Störbedingungen, welche das Testergebnis beeinflussen können.

Kontrollfragen

- Welcher der folgenden Sätze ist richtig?
 - a) Das aristotelische Wissensideal ist das Ideal der beweisenden Wissenschaft; oder
 - b) Das aristotelische Wissensideal ist das Ideal der erklärenden Wissenschaft?
- Worauf zielt das aristotelische Wissensideal ab?
- Was dient ihm als Vorbild?
- Was versteht man unter einem Axiom?
- Begründen Sie, inwiefern auch die Theorien der Naturwissenschaften strukturelle Theorieanteile enthalten!
- Nennen Sie einige psychologische Theorien, die strukturelle Gesetzmäßigkeiten untersuchen!
- Wann und von wem wurde der erste Intelligenztest konstruiert?
- Erläutern Sie die praktische Bedeutung der Verlässlichkeit psychologischer Testergebnisse!
- Wie wird die Retest-Reliabilität bestimmt?
- Worin besteht der Unterschied zwischen Retest-Reliabilität und Paralleltest-Reliabilität?
- Was versteht man in der klassischen Testtheorie unter True-Score und Messfehler?
- Wie ist der Messfehler definiert?
- Wie lautet die Grundgleichung der klassischen Testtheorie?
- Wie lauten die Axiome von Gulliksen?
- Wie ist die Reliabilität eines Tests definiert, welche Zahlenwerte kann sie annehmen, und was beschreiben diese?
- Welche Anforderungen sind an ein brauchbares Axiomensystem zu stellen?
- Was versteht man unter einem Kalkül?
- Was versteht man unter einem Modell?
- Diskutieren Sie Hilberts Forderung "Man muss jederzeit an Stelle von ‚Punkten, Geraden und Ebenen‘, ‚Tische, Stühle, Bierseidel‘ sagen können"! Was will er damit sagen?
- Welche Arbeitsteilung ergibt sich daraus zwischen den Formalwissenschaften einerseits und den Erfahrungswissenschaften andererseits?
- Nennen Sie Gründe, warum die Konstruktion eines psychologischen Modells der klassischen Testtheorie unerlässlich ist!
- Wie lautet das von Novick konstruierte Modell der klassischen Testtheorie, und auf welchen Grundannahmen beruht es?
- Welche Konsequenzen ergeben sich daraus für das Verständnis des Messfehlers der klassischen Testtheorie? Welche Fehlerquellen, die die Testergebnisse eines Probanden verzerren können, werden durch die klassische Testtheorie erfasst, und für welche ist das nicht der Fall?

- Zeigen Sie, dass der erwartete Messfehler für eine feste V_p v gleich Null ist!
- Zeigen Sie, dass für eine feste V_p v die gesamte Maßzahlvarianz Fehlervarianz ist!
- Beweisen Sie die Axiome von Gulliksen!
- Begründen Sie, warum der Messfehler der klassischen Testtheorie ausschließlich Zufallsfehler beschreibt!
- Begründen Sie, warum der True-Score der klassischen Testtheorie nicht die Leistungsfähigkeit eines Probanden in dem jeweiligen Test beschreibt, sondern lediglich seine Leistungsfähigkeit unter den jeweiligen Versuchs- und Störbedingungen, welche das Testergebnis beeinflussen können!

Übungsaufgaben

1. Der True-Score eines Probanden beträgt $\tau_{vt} = 47.8$ Punkte. Die Testung des Probanden ist mit einem Messfehler von 3.2 Punkten behaftet. Wie viele Punkte hat der Proband in dem Test erzielt?

2.2

Das galileische Wissensideal

Szientismus oder gegenstandsangemessene Umsetzung galileischen Denkens?

Das galileische Wissensideal ist das Ideal der erfahrungswissenschaftlichen Erklärung. Zunächst am Vorbild der Newton'schen Mechanik orientiert, hat dieses Wissensideal seit den Anfangstagen der wissenschaftlichen Psychologie wesentliche Präzisierungen und Erweiterungen erfahren. Diese sind einerseits den Grundlagenproblemen der noch jungen Human- und Sozialwissenschaften und andererseits der Entstehung der Wissenschaftstheorie als einer eigenen Disziplin geschuldet.

In Absetzung vom Programm des sogenannten Szientismus, wonach die Ideale und Begründungsverfahren der exakten Wissenschaften (insbesondere der harten Naturwissenschaften) auf die Theorienbildung in den Geistes- und Sozialwissenschaften übertragen werden sollen (Mittelstraß, 1996), sehen wir uns dabei in der Tradition von Kurt Lewin, der im Februar 1930 vor der Gesellschaft für empirische Philosophie in Berlin einen richtungsweisenden Vortrag gehalten hat, der später unter dem Titel *Der Übergang von der aristotelischen zur galileischen Denkweise in Biologie und Psychologie* in Band I der Zeitschrift *Erkenntnis* erschienen ist (Lewin, 1931).

Lewins Ziel war es dabei erklärter Weise nicht,

"... aus der Geschichte der Physik deduktiv zu schließen, was die Biologie tun 'soll'" (S. 423),

sondern im Gegenteil, er distanziert sich ausdrücklich von der Meinung,

"... dass es letzten Endes nur eine einzige empirische Wissenschaft, die Physik, gibt, auf die alle übrigen zurückgehen" (a.a.O.).

Gleichwohl ist sein Text als Plädoyer für eine moderne, erfahrungswissenschaftliche Auffassung von Psychologie zu lesen. Die Erfahrungswissenschaftlichkeit der Psychologie habe jedoch nicht aus einer Imitation der Physik, sondern aus einer gegen-

standsgemessenen Umsetzung des galileischen Denkens zu resultieren.

Eine dem Gegenstand der Psychologie angemessene Umsetzung des galileischen Denkens wird dabei möglich, wenn wir den Unterschied zwischen aristotelischem und galileischem Denken im Anschluss an Hoyningen-Huene (1989b) vor allem in zwei Punkten sehen:

Wesentliche Bestimmungsmerkmale des galileischen Denkens

- Die klassische (aristotelische) Naturwissenschaft zielt auf die Erkenntnis von Allgemeinem. Dieses Allgemeine sind einstellige Prädikate, die die wesentlichen Eigenschaften der Elemente einer bestimmten Klasse von Naturdingen angeben.
- Die neuzeitliche (galileische) Naturwissenschaft strebt zwar ebenfalls nach der Erkenntnis von Allgemeinem. Dieses Allgemeine sind jedoch Relationen, die – etwa im Falle der Physik – die naturgesetzlichen Zusammenhänge zwischen den Elementen bestimmter Klassen von materiellen Dingen angeben.

Geht man dabei – im Unterschied zum Szientismus – davon aus, dass es die Erklärung von Sachverhalten durch Herstellung von Relationen zwischen verschiedenen Klassen von Gegenständen sowie der empirische Gehalt dieser Erklärungen sind, die das galileische Denken ausmachen, nicht jedoch die (logische) Form der Sätze, durch welche diese Relationen beschrieben werden (Kempf, 1999), so lassen sich über die deduktiv-nomologische und die induktiv-statistische Erklärung hinaus auch das intentionale und das narrative Erklärungsmodell unter das galileische Denken subsumieren.

Kontrollfragen

- Welcher der folgenden Sätze ist richtig?
 - a) Das galileische Wissensideal ist das Ideal der beweisenden Wissenschaft; oder
 - b) Das galileische Wissensideal ist das Ideal der erklärenden Wissenschaft.
- Stellen Sie Lewins Wissenschaftsauffassung der des Szientismus gegenüber!
- Erläutern Sie, worin nach Hoyningen-Huene der Unterschied zwischen der klassischen (aristotelischen) und der neuzeitlichen (galileischen) Naturwissenschaft besteht!
- Erläutern Sie, unter welchem Verständnis des galileischen Denkens sich auch die intentionale und narrative Erklärung darunter subsumieren lassen!

2.3

Das deduktiv-nomologische Erklärungsmodell

Das Vorbild der
Newton'schen
Mechanik

Das deduktiv-nomologische Erklärungsmodell entspricht der Frage nach den Ursachen, als deren Wirkung der zu erklärende Sachverhalt eingetreten ist (vgl. Hempel, 1965). Es folgt damit dem Vorbild der Newton'schen Mechanik, welche die Bewegung von Körpern im dreidimensionalen euklidischen Raum erklärt, indem sie diese aus Naturgesetzen (gr. *nomos* = Gesetz) und sogenannten Randbedingungen ableitet (= deduziert).¹

Die ersten beiden
Newton'schen
Gesetze

Die ersten beiden Newton'schen Gesetze

1. Newton'sches Gesetz (Trägheitsprinzip): Jeder Körper verharrt im Zustand der Ruhe oder gleichförmiger, geradliniger Bewegung, falls er nicht durch äußere Kräfte gezwungen wird, diesen Zustand zu verlassen.

2. Newton'sches Gesetz (Aktionsprinzip): Die Beschleunigung (a) eines Körpers ist proportional zu der Kraft (F), die auf ihn ausgeübt wird, und umgekehrt proportional zu seiner Masse (m).

Formal: $a = F/m$ oder $F = m \times a$

Das Schwerkraft-
Gesetz

Das Schwerkraft-Gesetz

Die Kraft, mit der die Erde einen Gegenstand der Masse m anzieht, ist zum Erdmittelpunkt hin gerichtet.

Ihr Betrag ist $F = m \times g$.

Die Flugbahn einer
Kanonenkugel

So kann man z.B. mittels des zweiten Newton'schen Gesetzes und des Schwerkraft-Gesetzes die Flugbahn einer Kanonenkugel

1. Das Wort ‚Randbedingungen‘ wird hier im Sinne der wissenschaftstheoretischen Terminologie verwendet. In physikalischer Sprechweise unterscheidet man – je nach Situation – zwischen Anfangs- und/oder Randbedingungen, wobei das Wort ‚Randbedingungen‘ z.B. bei einer schwingenden Trommelfembran gebraucht wird, die am Rand festgespannt ist, so dass die Geschwindigkeit der Schwingungsbewegung dort immer gleich Null ist.

erklären, indem man die Anfangsbedingungen (d.h. die Koordinaten des Startortes sowie die horizontale und vertikale Komponente der Anfangsgeschwindigkeit, mit der die Kugel das Kanonenrohr verlässt) berücksichtigt und die Schwerkraft in Rechnung stellt, die auf die Kugel während ihres Fluges wirkt.¹

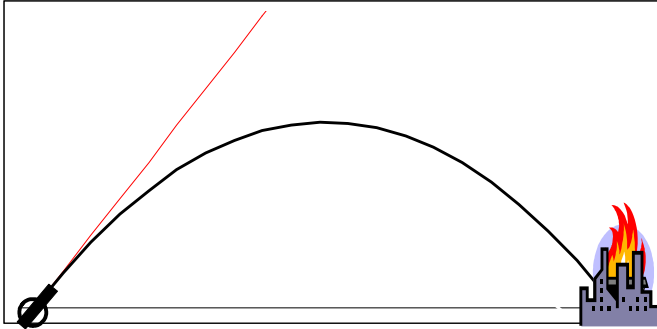


Abb. 2.3.1: Flugbahn einer Kanonenkugel

Von ihrer logischen Struktur her gesehen, folgt die deduktiv-nomologische Erklärung dem sog. HO-Schema (Hempel & Oppenheim, 1948), worin B(N) den zu erklärenden Sachverhalt (das Explanandum) und A(N) die Randbedingungen der Gesetzesaussage bezeichnet, die ihrerseits die Form einer universellen empirischen Subjunktion hat:

Das HO-Schema

A(N)	Randbedingung
$\forall x. A(x) \rightarrow B(x)$	Gesetzesaussage
<hr/>	
B(N)	Explanandum

Der Schluss von der Gesetzesaussage und den Randbedingungen auf das Explanandum folgt der logischen Schlussform des modus ponens, die wir bereits oben kennen gelernt haben:

$$(A(N) \wedge [A(N) \rightarrow B(N)]) \rightarrow B(N)$$

Nach diesem Erklärungsschema funktioniert z.B. auch das Weber'sche Gesetz, das – genau genommen – aus zwei Teilaussagen besteht:

Die beiden Teilaussagen des Weber'schen Gesetzes

1. Um die Flugbahn der Kanonenkugel auch tatsächlich berechnen zu können, müssen zuvor die Termini (Kraft, Masse, Beschleunigung) geklärt sein, die in den Gesetzen vorkommen. D.h. es muss festgelegt sein, wie man Kraft, Masse und Beschleunigung messen kann (vgl. z.B. Dransfeld, Kienle & Vonach, 1992). Erst auf dieser Grundlage kann man die Funktionsgleichung erstellen, welche die Flugbahn der Kanonenkugel beschreibt.

Die beiden Teilaussagen des Weber'schen Gesetzes

1. Teilaussage (absolute Unterschiedsschwelle): Zu jedem Standardreiz (S) gibt es einen Schwellenwert (ΔS), so dass der Unterschied (S-V) zwischen dem Standardreiz und einem Vergleichsreiz (V) genau dann wahrgenommen wird ($w(S-V)$), wenn die Reizdifferenz dem Betrage nach mindestens gleich ΔS ist. Formal:

$$(a) \quad \forall_{S,V}. |S-V| \geq \Delta S \rightarrow w(S-V)$$

$$(b) \quad \forall_{S,V}. |S-V| < \Delta S \rightarrow \neg w(S-V)$$

2. Teilaussage (relative Unterschiedsschwelle): Die absolute Unterschiedsschwelle (ΔS) ist proportional zum Standardreiz. Formal:

$$\forall_S. \Delta S = c \times S$$

Für alle Paare von Standardreiz und Vergleichsreiz gilt, dass der Unterschied zwischen den beiden Stimuli genau dann wahrgenommen wird, wenn er dem Betrag nach mindestens gleich $c \times S$ ist, andernfalls wird er nicht wahrgenommen. Die Randbedingungen der Gesetzesaussage sind einerseits die Größe des Standardreizes (S) und andererseits die Sinnesmodalität, durch welche die Größe der relativen Unterschiedsschwelle (c) bestimmt wird.

Eigenschaften von Gesetzesaussagen

Das Besondere an solchen Gesetzesaussagen besteht darin, dass sie universell gelten, obwohl es dafür keinen wie auch immer gearteten logischen oder sachlogischen Grund gibt. Dass das Gesetz gilt, zeigt allein die Erfahrung.

Das Zusammenwirken mehrerer Gesetzesaussagen

In der Regel wird man zur Erklärung eines Sachverhaltes (B) nicht nur eine Gesetzesaussage (G) und eine Randbedingung (R) heranziehen müssen, sondern mehrere Gesetzesaussagen mit entsprechenden Randbedingungen, welche gleichzeitig wirken. Die deduktiv-nomologische Erklärung nimmt dann die folgende Form an:

$$\begin{array}{l} R_1, R_2, \dots, R_m \\ G_1, G_2, \dots, G_n \\ \hline B \end{array}$$

Das Zusammenwirken der verschiedenen Gesetze und Randbedingungen kann in Form einer mathematischen Funktionsgleichung dargestellt werden:

$$B = f(R_1, R_2, \dots, R_m).$$

Zum Beispiel errechnet sich die in Abb. 2.3.1 dargestellte Flugbahn einer Kanonenkugel nach

$$x = v_{0,x} t \quad \text{und} \quad z = v_{0,z} t - \frac{1}{2} g t^2.$$

Daraus ergibt sich – zeitunabhängig geschrieben – die Bahngleichung

$$z = (v_{0,z}/v_{0,x}) x - (g/2(v_{0,x})^2) x^2.$$

Das ist die Gleichung einer Parabel.¹ Darin sind:

- x ... die (waagrechte) x-Koordinate der fliegenden Kugel,
- z ... die (senkrechte) z-Koordinate der fliegenden Kugel,
- $v_{0,x}$... die Anfangsgeschwindigkeit in x-Richtung,
- $v_{0,z}$... die Anfangsgeschwindigkeit in z-Richtung,
- g ... die Schwerebeschleunigung (9,81 m/s²),
- t ... die Zeit seit dem Start.

Voraussetzung für solche Berechnungen ist natürlich, dass sowohl das Explanandum als auch die Randbedingungen gemessen werden können. Denn nur als numerische Variablen können sie in Form einer mathematischen Funktionsgleichung ausgedrückt werden. Sowohl die Mathematisierung der Wissenschaften als auch die Messung folgen einem ganz konkreten praktischen Zweck: Sie dienen dazu, Naturgesetze exakt darzustellen und die gleichzeitige Wirkung mehrerer Naturgesetze beschreiben zu können. Weder Messung noch Mathematisierung sind folglich ein Wert an sich, der die Wissenschaftlichkeit einer Disziplin ausmacht, sondern sie sind praktische Erfordernisse, die sich erst aus einer bestimmten Form der erfahrungswissenschaftlichen Erklärung ergeben: aus der deduktiv-nomologischen und – wie wir später sehen werden – auch aus der induktiv-statistischen Erklärung.

Messung und
Mathematisierung
der Wissenschaften

Auch in der Psychologie müssen wir ggf. mit dem gleichzeitigen Wirken mehrerer Gesetze rechnen, wie der berühmte Asch-Ver-

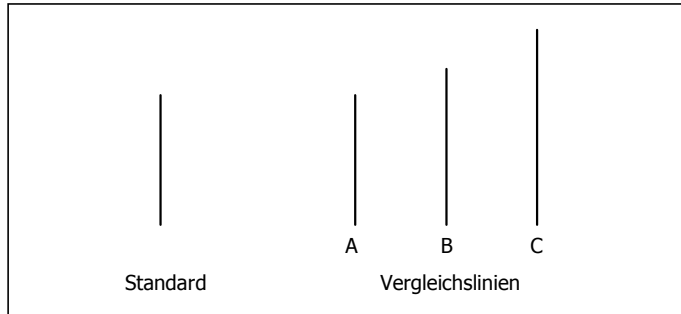
Der Asch-Versuch

1. Berechnungen dieser Art stellen natürlich stets Idealisierungen dar. So haben wir im konkreten Fall z.B. die Luftreibung vernachlässigt. Während man für einen mit der Hand geworfenen Ball noch gut mit der Schwerkraft als einziger Kraft auskommt, gilt dies schon nicht mehr für Tennisbälle oder weit geschossene Fußbälle. Moderne Ballistik ist noch weit komplizierter: nicht nur der Widerstand ist zu berücksichtigen, auch dynamischer Auf- und Quertrieb in der Luft, Kreiselkräfte etc.

sich zeigt, wonach die Unterschiedswahrnehmung durch Gruppendruck verändert werden kann.

Salomon Asch (1951, 1956) konfrontierte seine Versuchspersonen in einer Gruppensituation mit der Aufgabe, von drei Vergleichslinien jene zu benennen, die mit einer Standardlinie übereinstimmt (vgl. Abb. 2.3.2).

Abb. 2.3.2: Beispiel der Stimuli im Asch-Versuch



Das Stimulusmaterial war sehr eindeutig. Wenn Versuchspersonen einzeln befragt wurden, welche Vergleichslinie mit der Standardlinie übereinstimmt, ging die Zahl der Fehler gegen Null. Asch konfrontierte seine Versuchspersonen jedoch unter sozialem Einfluss mit dem Stimulusmaterial. So nahmen z.B. sieben Personen an dem Experiment teil, von denen nur eine, die sechste, eine echte Versuchsperson war. Die anderen waren Versuchsleitergehilfen, die instruiert waren, bei vorher festgelegten Stimuli übereinstimmend Falschurteile abzugeben und eine falsche Linie als gleichlang mit der Standardlinie zu bezeichnen. Dieses Arrangement führte dazu, dass die Probanden in etwa 30% der Fälle ebenfalls ein falsches Urteil abgaben.

Wollte man den Asch-Versuch mit Stimulusmaterial im Schwellenbereich wiederholen, so müssten die Wirkungen des Weber'schen Gesetzes und des Gruppendrucks gleichzeitig in Rechnung gestellt werden.

Kontrollfragen

- Wonach fragt das deduktiv-nomologische Erklärungsmodell?
- Was dient ihm als Vorbild?
- Welche logische Struktur hat es?
- Wie lauten die ersten beiden Newton'schen Gesetze?

- Wie lautet das Schwerkraft-Gesetz?
- Wie lautet das HO-Schema, und was stellt es dar?
- Mittels welcher logischen Schlussform kann von den Randbedingungen und der Gesetzesaussage auf den zu erklärenden Sachverhalt geschlossen werden?
- Geben Sie ein Beispiel für eine deduktiv-nomologische Erklärung in der Psychologie!
- Welche Art von Sachverhalten wird durch das Weber'sche Gesetz erklärt?
- Aus welchen beiden Teilaussagen besteht das Weber'sche Gesetz?
- Wie lauten die Randbedingungen des Weber'schen Gesetzes?
- Um welche Art von Aussagen handelt es sich bei den Gesetzesaussagen?
 - a) Partikuläre Aussagen,
 - b) universelle Aussagen,
 - c) sachlogische Aussagen oder
 - d) empirische Aussagen?
- Welches sind die wesentlichen Eigenschaften von Gesetzesaussagen?
- Wie kann man das Zusammenwirken mehrerer Gesetze darstellen? Geben Sie ein Beispiel!
- Welchem praktischen Zweck dient die Mathematisierung der (Natur-) Wissenschaften? Geben Sie ein Beispiel aus der Physik!
- Beschreiben Sie die Versuchsanordnung und die Ergebnisse des Asch-Versuches!

2.4

Messung

Messung als
Abbildung empirischer
auf numerische
Relationen

Unter einer Messung versteht man die Zuordnung von (Maß)Zahlen zu Messobjekten.

$$A \rightarrow x_A$$

Diese Zuordnung ist nicht Selbstzweck, sondern geschieht so, dass bestimmte empirisch feststellbare Relationen zwischen den Messobjekten auf entsprechende numerische Relationen zwischen den Maßzahlen abgebildet werden. Sind z.B. zwei Messobjekte A und B in einer bestimmten Hinsicht gleich ($A =^{\circ} B$), so erhalten sie auch gleiche Messwerte, so dass $x_A = x_B$.

$$A =^{\circ} B \rightarrow x_A = x_B$$

Der Sinn besteht darin, dass man von den Maßzahlen und den zwischen ihnen bestehenden Relationen jederzeit wieder auf die entsprechenden empirischen Relationen zwischen den Messobjekten zurückschließen können soll. Die Abbildung der empirischen Relationen auf die numerischen Relationen muss also homomorph sein, d.h. in beiden Richtungen gelesen werden können, z.B.

$$A =^{\circ} B \leftrightarrow x_A = x_B.$$

2.4.1 Qualitative und quantitative Variablen

Grundsätzlich kann bei den Messgrößen zwischen qualitativen und quantitativen Variablen unterschieden werden.

Qualitative Variablen

Qualitative Variablen sind z.B. das Geschlecht (männlich, weiblich), Blutgruppen (A, B, AB und 0) oder klinische Syndrome (die durch bestimmte Kombinationen von Krankheitssymptomen definiert sind).

Auch die Wahrnehmung des Unterschieds zwischen Standardreiz und Vergleichsreiz im Weber'schen Gesetz ist eine qualitative Va-

riable mit den beiden Ausprägungen $w(S-V)$ und $-w(S-V)$. Die Größe der Stimuli ist dagegen eine quantitative Variable, die je nach Art der Stimuli (Gewichte, Länge von Linien etc.) z.B. in Gramm oder in Zentimetern gemessen wird.

Auch qualitative Variablen können gemessen werden, indem man den verschiedenen Ausprägungen der Variable verschiedene Zahlen zuordnet. Man spricht dann von einer Nominalskala (welche die Messobjekte lediglich mit Zahlen benennt), z.B. Nominalskalen

$$\begin{aligned} w(S-V) &\rightarrow 1 \\ -w(S-V) &\rightarrow 0 \end{aligned}$$

oder

$$\begin{aligned} \text{weiblich} &\rightarrow 1 \\ \text{männlich} &\rightarrow 0 \end{aligned}$$

Welche Zahlen man zur Benennung der verschiedenen Qualitäten verwendet, ist völlig gleichgültig; wir könnten auch z.B. die Zahlen 87 und 13 als Maßzahlen verwenden. Wichtig ist nur, dass die beobachtete Information über die Messobjekte erhalten bleibt, d.h. aus den Maßzahlen ablesbar ist.

- Wenn $A = B$ (beide männlich oder beide weiblich), dann müssen sie auch dieselbe Maßzahl erhalten, so dass $x_A = x_B$.
- Wenn $A \neq B$ (verschiedenes Geschlecht), dann müssen sie auch verschiedene Maßzahlen erhalten, so dass $x_A \neq x_B$.

Unterscheidet eine Nominalskala nur zwei Skalenwerte, so spricht man von einer dichotomen Variable; unterscheidet sie mehr als zwei Skalenwerte, so spricht von einer polytomen Variable, z.B. Dichotome und polytome Variablen

$$\begin{aligned} \text{Blutgruppe A} &\rightarrow 1 \\ \text{Blutgruppe B} &\rightarrow 2 \\ \text{Blutgruppe AB} &\rightarrow 3 \\ \text{Blutgruppe 0} &\rightarrow 4 \end{aligned}$$

In der Praxis verwendet man als Skalenwerte meist die Zahlen 0,1,2,3... (das ist aber nicht zwingend). Auch auf die Reihenfolge kommt es nicht an. Eine dichotome Variable mit den Skalenwerten Null und Eins nennt man auch binär (weil die Ziffern 0 und 1 die Ziffern des binären Zahlensystems sind – im Unterschied zum dezimalen Zahlensystem, das die Ziffern 0, 1, 2, 3, 4, 5, 6, 7, 8 und 9 kennt).

Quantitative Variablen

Von einer quantitativen Variable spricht man dann, wenn die Zahlenwerte zumindest eine Rangordnung zwischen den Messobjekten abbilden. Voraussetzung dafür ist, dass wir eine empirische Ordnungsrelation zwischen den Messobjekten (z.B. $>^\circ$) beobachten können. Letzteres ist jedoch keineswegs trivial, da (zweifellose) empirische Relationen bestimmte Eigenschaften erfüllen müssen, damit sie eine Ordnungsrelation sind und auf die entsprechende numerische Ordnungsrelation (z.B. $>$) abgebildet werden können. Dazu müssen sie dieselben Eigenschaften haben wie die numerische Ordnungsrelation auch: Irreflexivität, Asymmetrie und Transitivität.

Soweit wir in der Psychologie auf physikalische Messungen objektseitig definierter Größen zurückgreifen können, braucht uns dies nicht weiter zu beunruhigen: z.B. bei der Messung von Standardreiz und Vergleichsreiz oder bei der Messung von Reaktionszeiten. Die prototheoretischen Probleme, die mit der Messung dieser Variablen verbunden sind, sind Gegenstand der Physik, und wir können auf funktionierende physikalische Messinstrumente zurückgreifen.

Der Nachweis einer empirischen Gesetzmäßigkeit als Voraussetzung der Messung

Anders sieht es dagegen mit der Messung psychologischer Größen aus, wie z.B. der Intelligenz oder von Einstellungen oder auch von subjektseitig definierten Größen, wie z.B. der subjektiven Reizstärke. Hier setzt die Messung dieser Größen den Nachweis einer empirischen Gesetzmäßigkeit voraus, den die Psychologie selbst erst erbringen muss.

Um z.B. rassistische Einstellungen messen zu können, genügt es nicht, den Probanden bestimmte Einstellungsfragen vorzulegen, wie z.B.

- "Schwarze sind weniger intelligent als Weiße",
- "Es würde mich stören, wenn ein Schwarzer in meine Nachbarschaft zieht", oder
- "Ich wäre nicht damit einverstanden, wenn meine Tochter einen Schwarzen heiratet".

Es genügt auch nicht, einen Probanden als umso rassistischer zu bezeichnen, je mehr dieser Fragen er zustimmend beantwortet. Die Anzahl der "positiven" Antworten liefert zwar eine (mehr oder minder angemessene) Quantifizierung des Ergebnisses der Befragung. Diese beschreibt jedoch nur die manifesten Antworten des Probanden, die sich vom Standpunkt eines außenstehenden

Beobachters (also objektseitig) feststellen lassen und nicht die latente (d.h. subjektseitig definierte) rassistische Einstellung des Probanden, für welche die Fragebogenantworten als Indikator dienen sollen. Dabei ist es noch nicht einmal ausgemacht, dass die Anzahl der positiven Antworten eine brauchbare Beschreibung des Befragungsergebnisses darstellen. Es wäre durchaus denkbar, dass manche Antworten stärker zu gewichten sind als andere. Oder es wäre auch denkbar, dass es so etwas wie eine Rassismus-Dimension, auf der sich die Probanden zwischen den Polen *rassistisch* und *nicht rassistisch* anordnen lassen, gar nicht gibt. Vielleicht gibt es stattdessen verschiedene Muster (Syndrome) von qualitativ verschiedenen rassistischen Einstellungen wie z.B.:

- Die einen sind überzeugt, dass Schwarze weniger intelligent sind und wollen deshalb nicht, dass ihre Tochter einen Schwarzen heiratet. Ansonsten haben sie aber nichts gegen Schwarze und würden sich auch nicht daran stören, wenn ein Schwarzer in ihre Nachbarschaft zieht.
- Andere fürchten im Gegenteil die mögliche intellektuelle Überlegenheit der Schwarzen, sind deshalb umso mehr um die Aufrechterhaltung der weißen Vorherrschaft besorgt und wollen daher Schwarze weder in ihrer Nachbarschaft noch als Schwiegersohn akzeptieren.

2.4.2 Die Guttman-Skala

Eine erste Antwort auf die Frage, welche empirischen Voraussetzungen erfüllt sein müssen, damit die Probanden aufgrund (binärer) Fragebogenantworten auf einer latenten Einstellungs-Dimension angeordnet werden können, wurde von Louis Guttman (1950) gegeben. Wie wir weiter unten sehen werden, stellt die sog. Guttman-Skala jedoch nur eine hinreichende, aber keine notwendige Bedingung dafür dar.

Sinngemäß geht Guttman davon aus, dass die Antworten (x) auf eine Rassismusfrage (i) nur dann etwas darüber aussagen, wie sich zwei Probanden bezüglich ihres Rassismus voneinander unterscheiden, wenn sie von dem einen Probanden (v) positiv ($x_{vi} = 1$) und von dem anderen (w) negativ ($x_{wi} = 0$) beantwortet wird. Ist dies der Fall, so gilt per definitionem jener Proband als

Definition einer Ordnungsrelation auf Grundlage der Fragebogenantworten

rassistischer ($v >^{\circ} w$), der die entsprechende Rassismus-Frage bejaht hat. Formal:

$$([x_{vi} = 1] \wedge [x_{wi} = 0]) \Rightarrow v >^{\circ} w$$

Somit definiert jede der Fragen eine Ordnungsrelation zwischen je zwei Probanden, und die Probanden können aufgrund ihrer Antworten genau dann auf einer Rassismus-Dimension angeordnet werden, wenn diese Relationen einander nicht widersprechen. D.h. es darf nicht vorkommen, dass ein Proband aufgrund der Antworten auf eine Frage als rassistischer erscheint als ein anderer Proband und aufgrund der Antworten auf eine andere Frage als weniger rassistisch.

Die nachzuweisende empirische Gesetzmäßigkeit

Diese Forderung, die sich unmittelbar aus der Asymmetrie von Ordnungsrelationen ergibt, hat zur Folge, dass man die Fragen (Items) so anordnen können muss, dass (z.B. bei $k=3$ Items) nur die folgenden Antwortmuster vorkommen:

Tab. 2.4.1: Zulässige Antwortmuster einer Guttman-Skala mit $k=3$ Items

0	0	0	0
1	0	0	1
1	1	0	2
1	1	1	3
0	1	2	

Damit werden gleichzeitig die Probanden (von oben nach unten) in eine Rangordnung bezüglich ihres Rassismus und die Fragen (von links nach rechts) in eine Rangordnung bezüglich ihrer Schwierigkeit gebracht. Je mehr Fragen ein Proband positiv beantwortet, desto rassistischer ist er; und je stärker der Rassismus eines Probanden ausgeprägt sein muss, damit er eine Frage positiv beantwortet, desto schwieriger ist die Frage. Die entsprechenden Rangplätze der Probanden stehen in der rechten Randspalte, die der Fragen in der unteren Randzeile der Tab. 2.4.1.

Erst die empirische Gesetzmäßigkeit, wonach von den 2^k möglichen Antwortmustern tatsächlich nur die oben dargestellten $k+1$ Muster vorkommen, und nicht etwa ein Muster der Form

$$1 \quad 0 \quad 1$$

erlaubt es somit, die Probanden auf einer quantitativen Rassismus-Dimension Θ anzuordnen und damit den Rassismus (ordinal) zu messen.

Verwenden wir o.B.d.A.¹ die nichtnegativen ganzen Zahlen als Skalenwerte, so lässt sich diese Gesetzmäßigkeit in der folgenden Formel ausdrücken, worin θ_v den Rangplatz des Probanden (v) und δ_i den Rangplatz der Frage (i) bezeichnet:

$$x_{vi} = \begin{cases} 1 & \text{wenn } \theta_v > \delta_i \\ 0 & \text{wenn } \theta_v \leq \delta_i \end{cases} \quad [3]$$

Der Rassismus der Probanden und die Schwierigkeit der Rassismus-Fragen werden auf derselben Rangskala gemessen.

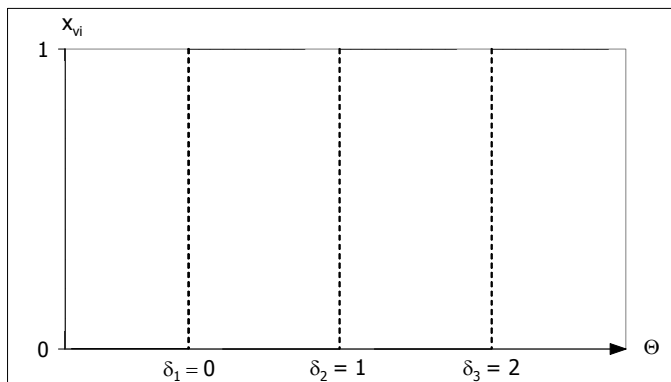


Abb. 2.4.1: Guttman-Skala mit 3 Items. Nichtnegative ganze Zahlen als Skalenwerte

Indem wir die Rangplätze mit den nichtnegativen ganzen Zahlen durchnummeriert haben, ist der Rangplatz eines Probanden auf der Rassismus-Skala gleich der Anzahl der von ihm gegebenen positiven Antworten. Im übrigen ist die Wahl der Skalenpunkte jedoch beliebig. Zu beachten ist lediglich,

- dass alle Probanden, welche dieselbe Anzahl von Fragen positiv beantwortet haben, denselben Rangplatz erhalten,
- dass dieser Rangplatz höher ist als der Rangplatz der schwierigsten Frage, welche die Probanden positiv beantwortet haben,
- jedoch höchstens so hoch wie der Rangplatz der leichtesten Frage, die sie negativ beantwortet haben.

Die gleichzeitige Messung des Rassismus der Vpn und der Schwierigkeit der Fragen auf derselben Ordinalskala

1. Die Abkürzung o.B.d.A. bedeutet "ohne Beschränkung der Allgemeinheit".

Eine metrische Information ist in der Guttman-Skala jedoch nicht enthalten: Die Abstände zwischen den Skalenwerten zweier Fragen sagen nichts darüber aus, wie stark sich die Fragen in ihrer Schwierigkeit unterscheiden, und die Abstände zwischen den Skalenwerten zweier Probanden sagen nichts darüber aus, wie sehr sie sich in ihrem Rassismus voneinander unterscheiden.

Genau so gut hätten wir die Skala auch so normieren können, dass die Schwierigkeiten der Fragen z.B. mit 1, 1.5 und 2.5 bezeichnet werden.

2.4.3 Skalentypen

In der Praxis gibt man sich mit bloßen Rangordnungen (Rang- oder Ordinalskalen) nicht gerne zufrieden. So werden wir z.B. die Entfernungen zu Urlaubszielen lieber in Kilometern messen als bloß auf einer Rangskala. Das Metermaß liefert uns nämlich mehr Informationen: Bilden wir nur eine Rangordnung, dann wissen wir lediglich, dass es z.B. von Konstanz an den Como-See näher ist als ans Mittelmeer. Geben wir die Entfernung dagegen in Kilometern an, dann wissen wir, dass wir ans Mittelmeer (von Konstanz aus) fast dreimal so weit fahren müssen.

Dass wir letzteres behaupten können, liegt daran, dass das Metermaß die Entfernung auf einer Verhältnisskala (oder Rationalskala) misst, die einen festen Nullpunkt hat und bei der gleiche Intervalle zwischen je zwei Messwerten auch gleiche Unterschiede zwischen den entsprechenden Messobjekten abbilden. Nur die Maßeinheit ist dabei willkürlich gewählt.

Einen Überblick über die wichtigsten Skalentypen und ihre Eigenschaften gibt Tab. 2.4.2.

Tab. 2.4.2: Überblick über die wichtigsten Skalentypen und ihre Eigenschaften

Skalentyp	Bildet ab:	Beispiel:
Nominalskala	Gleichheit oder Verschiedenheit der Messobjekte	Blutgruppen
	$A =^{\circ} B$ oder $A \neq^{\circ} B$	"A und B haben dieselbe Blutgruppe" oder "A und B haben verschiedene Blutgruppen"

Skalentyp	Bildet ab:	Beispiel:
Ordinalskala	Zusätzlich zur Nominalskala: Rangordnung der Messobjekte/besitzt keine Maßeinheit und keinen Nullpunkt	Sozioökonomischer Status (Unterschicht, Mittelschicht und Oberschicht)
	$A >^{\circ} B$	"A hat einen höheren sozioökonomischen Status als B"
Intervallskala	Zusätzlich zur Ordinalskala: Verhältnis der Unterschiede zwischen je zwei Messobjekten zueinander/Maßeinheit und Nullpunkt sind beliebig	Temperaturmessung in Celsius-Graden
	Der Unterschied zwischen A und B ist $(x_A - x_B)/(x_A - x_C)$ mal so groß wie der Unterschied zwischen A und C	Der Temperaturunterschied zwischen 2° und 6° Celsius ist doppelt so groß wie der zwischen 2° und 4°
Differenzskala	Zusätzlich zur Intervallskala: Unterschied zwischen je zwei Messobjekten/Feste Maßeinheit/Nullpunkt ist beliebig	Psychometrie: logistisches Testmodell $\ln\{p_{vi}/(1 - p_{vi})\} = \theta_v - \delta_i$ p_{vi} = Lösungswahrscheinlichkeit der Vp v bei Aufgabe i θ_v = Leistungsfähigkeit der Vp δ_i = Schwierigkeit der Aufgabe i
	Der Unterschied zwischen A und B beträgt $x_A - x_B$	Zwei Vpn A und B unterscheiden sich in ihrer Leistungsfähigkeit um 0,7 Skalenpunkte
Rationalskala	Zusätzlich zur Intervallskala: Verhältnis je zweier Messobjekte zueinander/Fester Nullpunkt/Maßeinheit ist beliebig	Längenmessung
	Das Verhältnis von A zu B beträgt x_A/x_B	Die Entfernung von 100 km ist halb so weit wie die Entfernung von 200 km

Skalentyp	Bildet ab:	Beispiel:
Absolutskala	Zusätzlich zur Intervallskala: Unterschied zwischen je zwei Messobjekten <i>und</i> Verhältnis je zweier Messobjekte zueinander sowie der Zahlenwert selbst/ Fester Nullpunkt und feste Maßeinheit	Wahrscheinlichkeiten und relative Häufigkeiten (als empirische Schätzung der Wahrscheinlichkeit)
	A hat den Wert x_A , der Unterschied zwischen A und B beträgt $x_A - x_B$ <i>und</i> das Verhältnis von A zu B beträgt x_A/x_B	Die Wahrscheinlichkeit, dass eine zufällig herausgegriffene V_p bei den letzten Wahlen die Partei x gewählt hat, ist 0.10 (10%). Dies bedeutet einen Zuwachs von +0.05 (5%) gegenüber den vorangegangenen Wahlen, womit sich der Wähleranteil dieser Partei verdoppelt hat.

2.4.4 Allgemeine und angewandte Messtheorie

Das Repräsentationsproblem

Um eine höhere Skala konstruieren zu können, reicht es freilich nicht aus, über empirische Ordnungsrelationen zwischen je zwei Messobjekten zu verfügen. Man braucht zusätzlich mehrstellige Ordnungsrelationen, die z.B. den empirischen Unterschied oder das empirische Verhältnis je zweier Messobjekte betreffen:

$$U(A,B) >^{\circ} U(C,D)$$

oder

$$V(A,B) >^{\circ} V(C,D)$$

Unter dem Stichwort ‚Repräsentationsproblem‘ untersucht die allgemeine Messtheorie, welche Art von empirischen Relationen erforderlich bzw. ausreichend ist, um eine Skala von dieser oder jener Qualität zu konstruieren.

Einerseits geht es dabei um die Frage, welche Eigenschaften die empirischen Relationen erfüllen müssen, damit eine bestimmte Art von Skalen konstruiert werden kann. Andererseits geht es um die Frage, welche Art von Skalen konstruiert werden kann, wenn empirische Relationen bestimmte Eigenschaften aufweisen.

Ein grundlegendes Ergebnis der allgemeinen Messtheorie haben wir bereits kennen gelernt, indem wir festgestellt haben, dass zur Konstruktion einer Ordinalskala das Vorliegen einer zweistelligen empirischen Ordnungsrelation erforderlich ist, welche die Eigenschaften der Irreflexivität, Asymmetrie und Transitivität besitzt.

Im Unterschied zur allgemeinen Messtheorie befasst sich die angewandte Messtheorie mit dem Skalierungsproblem. Sie untersucht, ob gegebene empirische Relationen, auf deren Grundlage eine Skala konstruiert werden soll, die geforderten Eigenschaften auch tatsächlich erfüllen. Darauf aufbauend, konstruiert sie schließlich Messinstrumente, welche den Messobjekten Maßzahlen zuordnen, so dass die empirischen Relationen zwischen den Messobjekten auf die entsprechenden numerischen Relationen zwischen den Maßzahlen abgebildet werden.

Das
Skalierungsproblem

Die angewandte (psychologische) Messtheorie konstruiert damit psychologische Modelle der – von der allgemeinen Messtheorie entwickelten – (formalen) messtheoretischen Kalküle. Als ein erstes Beispiel dafür haben wir bereits oben die Guttman-Skala kennengelernt.

Unter dem Stichwort ‚Eindeutigkeitsproblem‘ untersucht die allgemeine Messtheorie die Skaleneigenschaft von Messungen. Dabei ist der Skalentyp einer Messung durch die Klasse der zulässigen Maßzahltransformationen definiert, d.h. dadurch, wie man die Maßzahlen verändern darf, ohne dass die Eigenschaften einer Skala dadurch verloren gehen.

Das
Eindeutigkeitsproblem

Eine Rationalskala haben wir z.B. dadurch beschrieben, dass sie einen festen Nullpunkt hat, und dass lediglich die Maßeinheit willkürlich gewählt wurde. Zulässig sind hier also nur solche Transformationen, welche zwar ggf. die Maßeinheit verändern, aber den Nullpunkt unverändert lassen. Das sind Transformationen der Art, dass wir die Messwerte mit einer Konstanten multiplizieren, die allerdings größer als Null sein muss:

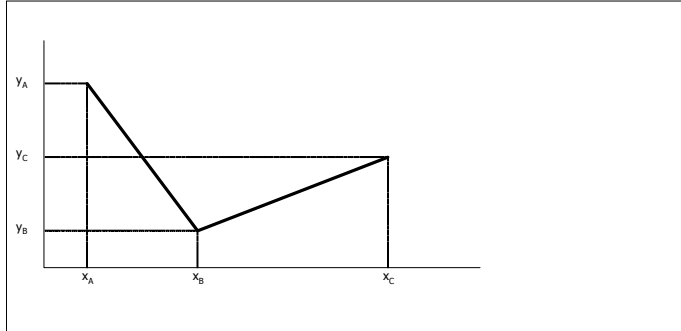
$$X \rightarrow Y: x \rightarrow y = cx, \quad \text{mit } c > 0$$

Ist die Konstante $c < 0$, so verkehrt sich die Skala in ihrer Bedeutung in ihr Gegenteil: Aus $x_A > x_B$ wird $y_A < y_B$. Hat die ursprüngliche Skala (X) z.B. die Länge gemessen, so misst die transformierte Skala (Y) nun die Kürze. Ist $c = 0$, so wird die Skala gänzlich zerstört, und allen Messobjekten wird dieselbe Maßzahl $y = 0$ zugewiesen.

Nominalskala

Je aussagekräftiger eine Skala ist, desto stärkeren Restriktionen unterliegt die Klasse der zulässigen Maßzahltransformationen. So ist z.B. bei Nominalskalen jede eineindeutige Maßzahltransformation zulässig, welche gleiche Maßzahlen auf gleiche und verschiedene Maßzahlen auf verschiedene abbildet (vgl. Abb. 2.4.2).

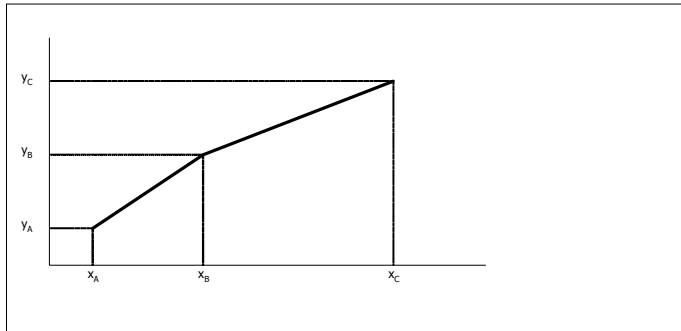
Abb. 2.4.2: Nominalskala:
Eineindeutige Maßzahl-
transformation



Ordinalskala

Bei der Ordinalskala sind dagegen nur noch streng monoton wachsende Transformationen zulässig, so dass die Größer-kleiner-Relation erhalten bleibt (vgl. Abb. 2.4.3).

Abb. 2.4.3: Ordinalskala:
Streng monoton wachsen-
de Maßzahltransformation



Intervallskala

Bei der Intervallskala sind nur noch positive lineare Transformationen der Form

$$X \rightarrow Y: x \rightarrow y = a + cx, \quad \text{mit } c > 0$$

zulässig, welche den Nullpunkt der Skala um den Betrag a verschieben und die Maßeinheit um den Faktor c vergrößern bzw. verkleinern (vgl. Abb. 2.4.4).

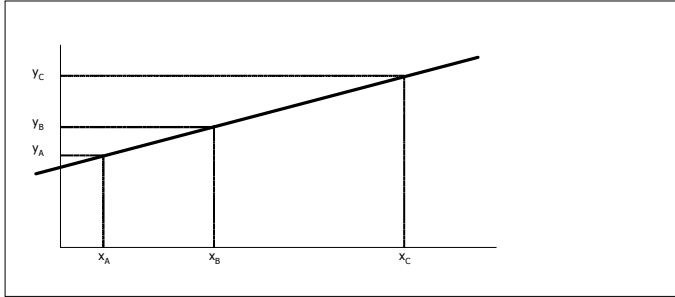


Abb. 2.4.4: Intervallskala:
Positive lineare Transformation

Bei der Differenzskala ist die lineare Transformation so eingeschränkt, dass $c = 1$. D.h. es ist nur noch die Addition einer Konstanten zulässig, welche den Nullpunkt der Skala verschiebt, während die Maßeinheit unverändert bleibt (vgl. Abb. 2.4.5).

Differenzskala

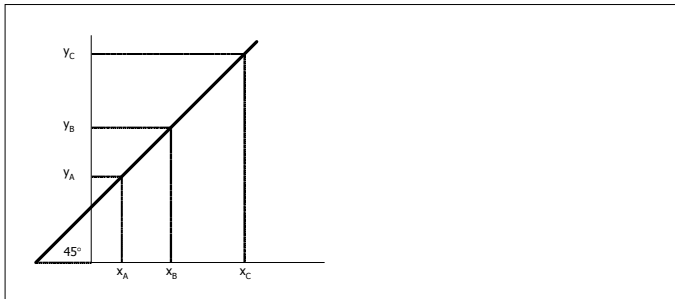


Abb. 2.4.5: Differenzskala:
Addition einer Konstanten

Bei der Rationalskala ist die lineare Transformation dagegen so eingeschränkt, dass $a = 0$. D.h. es ist nur noch die Multiplikation mit einer positiven Konstanten zulässig, welche die Maßeinheit verändert, während der Nullpunkt erhalten bleibt (vgl. Abb. 2.4.6).

Rationalskala

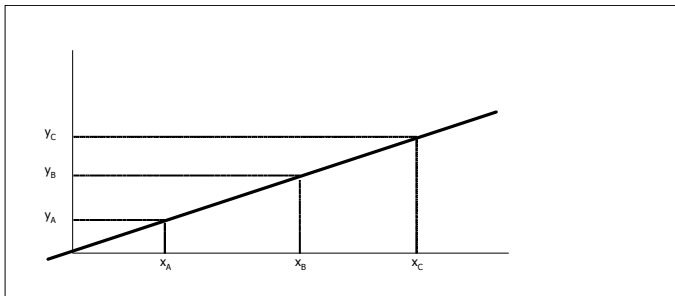
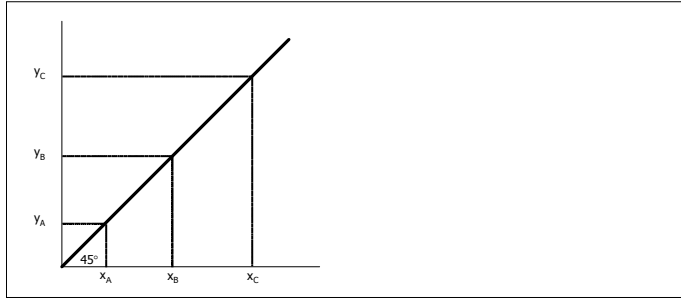


Abb. 2.4.6: Rationalskala:
Multiplikation mit einer positiven Konstanten

Absolutskala

Bei der Absolutskala ist keinerlei Transformation der Skalenwerte mehr zulässig bzw. nur die sog. identische Transformation, welche jeden Skalenwert in sich selbst überführt, so dass $Y = X$ (vgl. Abb. 2.4.7).

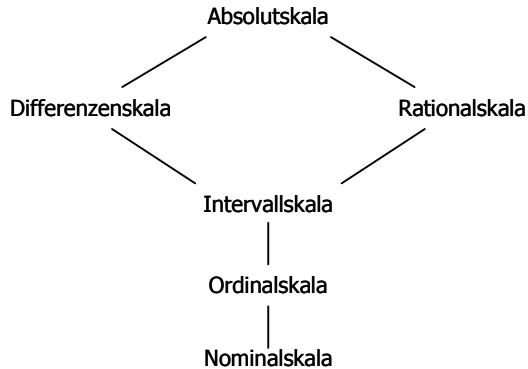
Abb. 2.4.7: Absolutskala:
Identische Transformation



Die Hierarchie der Skalentypen

Zwischen den Skalentypen besteht folglich eine Hierarchie, die in Abb. 2.4.8 dargestellt ist:

Abb. 2.4.8: Hierarchie der Skalentypen



Wird auf eine höhere Skala eine Transformation angewendet, die nur für einen niedrigeren Skalentyp zulässig ist, so wird dadurch die Skala nicht gänzlich zerstört. Der Skalentyp sinkt jedoch auf das Niveau der in der Hierarchie tiefer stehenden Skala ab. Man erhält also eine "schlechtere" Skala, die weniger aussagekräftig ist.

Das Bedeutsamkeitsproblem

Dies zu beachten ist wichtig, weil der Skalentyp einer Messung eng mit dem Bedeutsamkeitsproblem zusammenhängt. Bedeutsam – und damit sinnvoll interpretierbar – sind über Maßzahlen getroffene Aussagen nur dann, wenn ihr Wahrheitswert gegenüber den

zulässigen Maßzahltransformationen unveränderlich bleibt. Dadurch beschränkt der Skalentyp auch die Art der Rechenoperationen, die mit den Maßzahlen angestellt werden dürfen, und die Art der statistischen Methoden, welche (sinnvollerweise) auf die Daten angewandt werden können. Einen Überblick über die zulässigen Maßzahltransformationen, über bedeutsame Aussagen und zulässige statistische Verfahren gibt Tab. 2.4.3.

Skalentyp	Zulässige Maßzahltransformation	Bedeutsame Aussagen und zulässige statistische Verfahren
Nominalskala	Jede eindeutige Transformation, so dass $x_A = x_B \rightarrow y_A = y_B$ und $x_A \neq x_B \rightarrow y_A \neq y_B$	Gleichheit/Ungleichheit von Skalenwerten
		Statistische Verfahren für qualitative Variablen, z.B. Häufigkeitsverteilung, Modus, Kontingenztafeln, Pearson- χ^2 , log-lineares Modell
Ordinalskala	Jede streng monoton wachsende Transformation, so dass $x_A > x_B \rightarrow y_A > y_B$	Rangordnung der Skalenwerte
		Nonparametrische statistische Verfahren (z.B. Wilcoxon-Test), Median, Quartilabstand, Koeffizient τ nach Kendall
Intervallskala	Jede positive lineare Transformation der Art $y = a + cx$, mit $c > 0$	Vergleich der Summen und Differenzen von Skalenwerten
		Zusätzlich zu Ordinalskala: parametrische statistische Verfahren (t-Test, F-Test, lineares Modell etc.), Mittelwert, Standardabweichung, Produktmomentkorrelation

Tab. 2.4.3: Überblick über zulässige Maßzahltransformationen, bedeutsame Aussagen und zulässige statistische Verfahren

Skalentyp	Zulässige Maßzahltransformation	Bedeutsame Aussagen und zulässige statistische Verfahren
Differenzskala	Addition einer Konstanten, so dass $y = a + x$	Vergleich der Summen und Differenzen von Skalenwerten
		Siehe Intervallskala und ggf. weitere
Rationalskala	Multiplikation mit einer positiven Konstanten, so dass $y = cx$, mit $c > 0$	Vergleich der Summen, Differenzen, Produkte und Quotienten von Skalenwerten
		Siehe Intervallskala und ggf. weitere
Absolutskala	Keinerlei (bzw. nur die identische Transformation) $y = x$	Alle Rechenoperationen
		Siehe Intervallskala und ggf. weitere

Parametrische statistische Verfahren

Eine verbreitete Fehlerquelle empirischer psychologischer Untersuchungen stellt die routinemäßige Anwendung parametrischer statistischer Verfahren (z.B. Mittelwert, Standardabweichung, Produktmomentkorrelation) dar, die zur Standardausbildung in Statistik für Psychologen gehören. Da zu ihrer Berechnung die Rechenoperation der Summenbildung erforderlich ist, setzen diese Verfahren (mindestens) Intervallskalenqualität der Daten voraus. Die Messung psychologischer Variablen ist jedoch in der Regel nur nominal oder ordinal, und zwar selbst dann, wenn als Indikatoren der psychologischen Variablen auf Beobachtungsgrößen zurückgegriffen werden kann, die auf einer höheren Skala gemessen werden.

Manifeste und latente Variablen

So ist z.B. ‚Hunger‘ eine subjektseitig definierte (latente) Variable, für die sich verschiedene objektseitig definierte (manifeste) Indikatoren angeben lassen, die sich – wie z.B. die Dauer des Nahrungsentzuges – z.T. sogar auf einer Rationalskala messen lassen. Zwischen diesem Indikator und dem Hunger besteht ein streng monoton wachsender Zusammenhang: Je länger der Nahrungsentzug ist, desto größer ist der Hunger. Solange über die genaue Form dieses Zusammenhangs jedoch nichts bekannt ist,

können wir die Dauer des Nahrungsentzuges noch so genau messen, über den Hunger werden wir daraus nur ordinale Informationen beziehen.

Um ordinale Daten statistisch weiterzuverarbeiten, muss man sich nonparametrischer Verfahren bedienen, die allerdings mehr als bloße Grundkenntnisse in Statistik voraussetzen.

Nonparametrische statistische Verfahren

Die oft fälschliche Anwendung parametrischer Verfahren in der Psychologie wird zudem durch ein (selbst in Lehrbüchern, z.B. Huber, 1984, 87) verbreitetes Missverständnis gefördert, welches in die Zeit zurückreicht, als es in der Psychologie zwar schon Intelligenz- und andere Tests gab, die Messtheorie aber noch nicht so weit entwickelt war, dass man einen klaren Begriff davon hatte, was Messung denn eigentlich ist und wie der Skaleneigentyp einer Messung sichergestellt werden kann. Dieses Missverständnis besagt z.B., dass der Intelligenzquotient (IQ) auf einer Intervallskala gemessen werde, da er (näherungsweise) normalverteilt sei.

Skaleneigenschaften und Verteilungseigenschaften

Richtig daran ist zunächst, dass die Normalverteilung – ebenso wie Intervallskalen auch – gegenüber positiven linearen Transformationen invariant ist, d.h. dass die Verteilungsform der Normalverteilung erhalten bleibt, wenn man die Maßzahlen positiv linear transformiert.

Wurde eine Variable auf einer Intervallskala gemessen und ist diese Variable normalverteilt, dann bleibt die Normalverteilung der Variable unter allen zulässigen Maßzahltransformationen erhalten. Umgekehrt kann jedoch nicht von der Normalverteilung auf eine Intervallskala geschlossen werden, da Verteilungstyp und Skalentyp nichts miteinander zu tun haben (vgl. Fischer, 1968).

2.4.5 Manifeste und latente Variablen

Manifeste Variablen sind solche, die sich direkt beobachten lassen, wie z.B. die Testleistung in einem Intelligenztest. Latente Variablen können dagegen nicht direkt beobachtet werden, sondern sind aus beobachtbaren Variablen abgeleitet, z.B. die Intelligenz.

Definition

Um von manifesten Variablen auf latente Variablen schließen zu können, muss eine gesetzmäßige Beziehung angegeben werden, wie die zu erschließende, latente Variable mit den beobachteten, manifesten Indikatoren zusammenhängt.

Die Itemcharakteristikkfunktion

In der differentiellen Psychologie postuliert man z.B.:

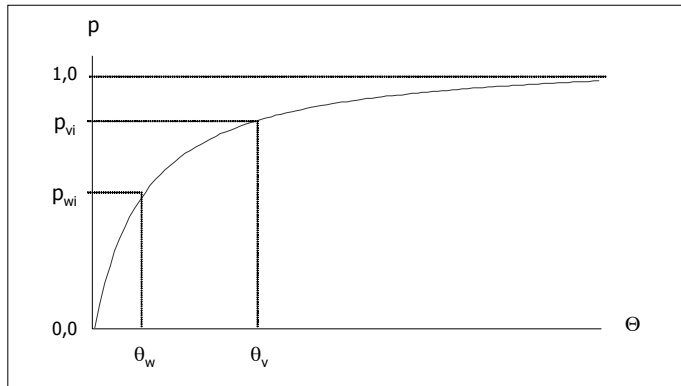
- Je intelligenter eine Vp (v) ist, desto eher wird sie Aufgaben (Items) einer bestimmten Art (I) lösen können.
- Und umgekehrt schließt man: Je wahrscheinlicher es ist, dass eine Vp eine Aufgabe ($i \in I$) lösen kann, desto intelligenter ist sie.

Die Wahrscheinlichkeit des manifesten Indikators ‚Lösung der Testaufgabe‘ (p_{vi}) kann also als streng monoton wachsende Funktion der latenten Dimension ‚Intelligenz‘ (Θ) dargestellt werden, wobei θ_v die Intelligenz einer bestimmten Vp (v) darstellt:

$$\forall_v \forall_{i \in I}: p_{vi} = f_i(\theta_v) \quad [4]$$

Diese Funktion wird in der Psychometrie als Itemcharakteristik bezeichnet.

Abb. 2.4.9: Beispiel einer Itemcharakteristik



Über die Form der Itemcharakteristik ist zunächst nur so viel vorausgesetzt, dass es sich um eine streng monoton wachsende Funktion handelt (vgl. Abb. 2.4.9), so dass von je zwei Probanden v und w der intelligentere ($\theta_v > \theta_w$) die größere Lösungswahrscheinlichkeit ($p_{vi} > p_{wi}$) hat und umgekehrt

$$\theta_v > \theta_w \Leftrightarrow p_{vi} > p_{wi} \quad [5]$$

Die Skaleneigenschaften der beteiligten Variablen

In Gleichung [4] haben wir es mit drei verschiedenen Arten von Variablen zu tun:

1. Mit manifesten Antwortvariablen X_{vi} . Diese beschreiben die Antwort, welche Vp v auf Item i gibt, und sind qualitative (dichotome) Variablen (gelöst/nicht gelöst), die auf einer Nominalskala gemessen werden:

$$x_{vi} = \begin{cases} 1 & \text{falls Item } i \text{ gelöst} \\ 0 & \text{falls Item } i \text{ nicht gelöst} \end{cases}$$

2. Mit Lösungswahrscheinlichkeiten $p_{vi} = p(X_{vi} = 1)$. Diese sind nicht direkt beobachtbar (also latent) und werden auf einer Absolutskala gemessen.
3. Mit der Intelligenz Θ , die ebenfalls latent ist und deren Skaleneigenschaft davon abhängt, welche Form die Itemcharakteristiken haben.

Genauer gesagt, wird die Skaleneigenschaft der latenten Dimension Θ wieder durch die Klasse der zulässigen Maßzahltransformationen bestimmt, welche die Form des funktionalen Zusammenhangs zwischen den Lösungswahrscheinlichkeiten p_{vi} und der latenten Dimension Θ nicht zerstören.

Haben die Itemcharakteristiken z.B. die Form der logistischen Funktion

Die logistische Funktion

$$p_{vi} = \frac{e^{(\theta_v - \delta_i)}}{1 + e^{(\theta_v - \delta_i)}} = f_i(\theta_v) \quad [6]$$

bzw. gleichwertig

$$\theta_v - \delta_i = \ln\left(\frac{p_{vi}}{1 - p_{vi}}\right) \quad [7]$$

dann wird die Schwierigkeit der Items (δ_i) auf derselben Differenzskala (Θ) gemessen wie die Intelligenz der Vpn (θ_v) (vgl. Abb. 2.4.10). Die Maßzahltransformation

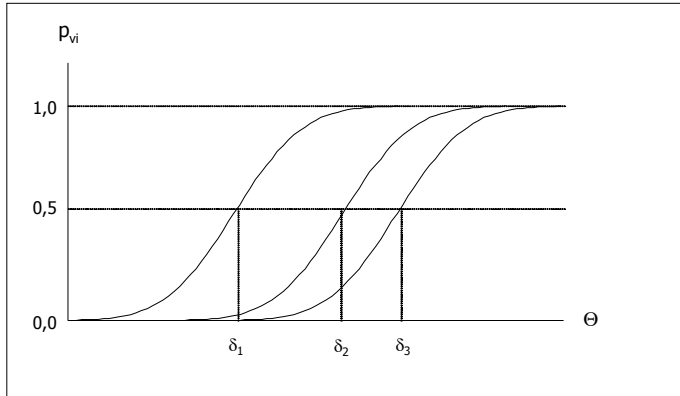
$$\begin{aligned} \Theta \rightarrow \Theta': \theta_v &\rightarrow \theta'_v &= a + \theta_v \\ &\delta_i &\rightarrow \delta'_i &= a + \delta_i \end{aligned}$$

liefert eine gleichwertige Skala (Θ') mit

$$\theta'_v - \delta'_i = \theta_v - \delta_i,$$

so dass die Funktionsgleichungen [6] und [7] erhalten bleiben, wenn man die Skalenwerte durch die transformierten Werte ersetzt. Die Skala ist somit bis auf eine additive Konstante bestimmt, und lediglich der Nullpunkt kann willkürlich festgesetzt werden.

Abb. 2.4.10: Logistische Itemcharakteristiken dreier Items mit unterschiedlicher Schwierigkeit (δ_i). Je größer die latente Fähigkeit (θ), desto größer die Lösungswahrscheinlichkeiten. Je schwieriger ein Item, desto fähiger muss die V_p sein, um eine bestimmte Erfolgswahrscheinlichkeit zu erreichen. Halten sich Itemschwierigkeit und Fähigkeit der V_p die Waage ($\delta_i = \theta_v$), so ist die Lösungswahrscheinlichkeit gleich $\frac{1}{2}$



Zusammenfassung

Zusammenfassend können wir festhalten: Wird eine theoretische (latente) Variable durch beobachtbare (manifeste) Indikatoren operationalisiert, so gehört die Messung der theoretischen Variable in der Regel nicht demselben Skalentyp an wie die Messung der manifesten Variable, durch welche sie operationalisiert wird. Die Messung der theoretischen Variable kann sowohl höhere (wie im hier diskutierten Beispiel) als auch geringere Skaleneigenschaften haben (wie im Beispiel der Operationalisierung des Hungers durch die Dauer des Nahrungszuges). Welchem Skalentyp die Messung der latenten Variable angehört, hängt von der Spezifikation des funktionalen Zusammenhanges der (latenten) Variable mit ihren (manifesten) Indikatoren sowie auch von der Skaleneigenschaft der Indikatoren ab.

Stochastische Testmodelle

Testtheoretische Modelle, welche die Itemantworten mittels der Itemcharakteristikfunktion auf eine quantitative latente Eigenschaft (engl. *trait*) der Probanden zurückführen, bezeichnet man in der Psychometrie als Latent-Trait-Modelle. Ist die latente Variable dagegen qualitativer (klassifikatorischer) Art, so spricht man von Latent-Class-Modellen. Latent-Trait- und Latent-Class-Modelle zusammenfassend, spricht man von Item-Response-Modellen und unterscheidet diese von der klassischen Testtheorie, welche die Testscores (z.B. die Anzahl der gelösten Items) zum Ausgangspunkt der Theorienbildung nimmt, deren Zustandekommen selbst jedoch nicht weiter problematisiert. Die Art der

statistischen Modellbildung (vgl. Kapitel 2.6.2) hervorhebend, spricht man von stochastischen oder probabilistischen Testmodellen. Lazarsfeld (1950, 1959), der Begründer der stochastischen Testtheorie, verwendete die Bezeichnung ‚Latent Structure Analysis‘, die sich aber nicht allgemein durchsetzen konnte.

2.4.6 Latent-Trait-Modelle¹

Wie wir bereits gesehen haben, ist über die Form der Itemcharakteristiken zunächst nur so viel bekannt, dass es sich dabei um streng monoton wachsende Funktionen handelt. Die dadurch definierte zweistellige empirische Ordnungsrelation [5] ist jedoch (allenfalls) für die Konstruktion einer Ordinalskala ausreichend, was sofort die Fragen aufwirft,

Die Spezifikation der Itemcharakteristiken

1. wie man die Itemcharakteristiken zu spezifizieren hat, um eine leistungsfähigere Skala zu ermöglichen, und
2. wie man überprüfen kann, ob die derart getroffenen Modellannahmen auch mit der empirischen Wirklichkeit übereinstimmen.

Insbesondere die letztgenannte Frage ist keineswegs trivial, denn zum einen kann noch nicht einmal a priori vorausgesetzt werden, dass die Items eines Tests tatsächlich alle dieselbe latente Eigenschaft der Probanden messen, und zum anderen muss die empirische Prüfung der Modellannahmen auch insofern problematisch erscheinen, als die latente Dimension ja überhaupt nur durch die in der Itemcharakteristik ausgedrückte Verknüpfung mit den manifesten Itemantworten definiert ist und nicht unabhängig von ihr gemessen werden kann.

Was die Spezifikation der Itemcharakteristiken betrifft, so wurden im Laufe der Geschichte verschiedene Vorschläge gemacht. Paul Lazarsfeld (1950, 1959, 1961; Lazarsfeld & Henry, 1968) experimentierte mit linearen Itemcharakteristiken, und Frederick Lord (1952, 1953a, 1953b; Lord & Novick, 1968) versuchte, die Itemcharakteristiken durch die Verteilungsfunktion der Normalverteilung darzustellen (Normal-Ogive-Modell). Die Annahme, dass die Itemcharakteristiken der in Gleichung [6] dargestellten logistischen Funktion folgen, geht auf den dänischen Statistiker Georg Rasch zurück und wird in der Literatur als Rasch-Modell

Das Rasch-Modell

1. Dieser Abschnitt kann ohne Verlust der Kontinuität übersprungen werden.

bezeichnet. Das Rasch-Modell und seine verschiedenen Varianten und Weiterentwicklungen (Rasch, 1960, 1961; Birnbaum, 1968; Micko 1970; Kempf, 1972; Fischer, 1974b; Andrich 1978a, b; Rost, 1990) können heute als die am besten fundierte Latent-Trait-Theorie gelten.

Das Skalierungsproblem

Um zu verstehen, warum sich gerade die logistische Funktion in der Psychometrie so erfolgreich durchsetzen konnte, müssen wir uns vergegenwärtigen, dass die Lösung des Skalierungsproblems im Kontext von Latent-Trait-Modellen nichts anderes bedeutet als die Schätzung unbekannter Parameter, welche die Wahrscheinlichkeitsverteilung der Antworten von n Vpn auf k Testitems beschreiben, so dass die Likelihood (Wahrscheinlichkeit) der Antwortmatrix $\mathbf{X} = ((x_{vi}))$ in der Form

$$L(\mathbf{X}) = \prod_{v=1}^n \text{prob}\{(x_{v1}, \dots, x_{vk}) | \theta_v\} \quad [8]$$

dargestellt werden kann. Als Funktion der Itemcharakteristiken $f_i(\theta_v)$ werden die Wahrscheinlichkeiten $\text{prob}\{(x_{v1}, \dots, x_{vk}) | \theta_v\}$ dabei außer von dem Personenparameter θ_v auch von k (ein- oder mehrdimensionalen) Itemparametern δ_i ($i = 1, \dots, k$) abhängen, welche die relevanten Eigenschaften der Testaufgaben (Schwierigkeit, Trennschärfe etc.) messen.

Statistische Schätzprobleme

Dies wirft erhebliche Schätzprobleme auf, da die herkömmlichen Methoden der Parameterschätzung, namentlich die Maximum-Likelihood-Methode, versagen, wenn die Anzahl der zu schätzenden Parameter nicht gegen einen festen Wert konvergiert, während die Anzahl der Beobachtungen gegen unendlich strebt. Gerade diese Voraussetzung wird durch die Latent-Trait-Modelle aber verletzt. Um die Verteilung der $n \times k$ Antworten von n Vpn auf k Testitems zu beschreiben, müssen (mindestens) $n+k$ Parameter geschätzt werden, und es gibt keine Möglichkeit, wie die Anzahl der Beobachtungen vergrößert werden könnte, ohne dass die Anzahl der zu schätzenden Parameter mit wächst. Einem Satz von Neyman & Scott (1948) zufolge liefert die Maximum-Likelihood-Methode in diesem Falle nicht einmal konsistente Schätzfunktionen.

Die Separierbarkeit der Modellparameter

Mit der Separierbarkeit der Modellparameter hat die logistische Funktion eine Eigenschaft, die sie in pragmatischer, statistischer und wissenschaftstheoretischer Hinsicht gegenüber allen anderen denkbaren Itemcharakteristiken auszeichnet und die zugleich

die eleganteste und mathematisch-statistisch zufriedenstellendste Lösung der skizzierten Schätzprobleme gewährleistet.

Grundlage der Separierung der Modellparameter ist die Existenz erschöpfender (suffizienter) Statistiken für die Personenparameter: Die logistische Itemcharakteristik ist notwendig und hinreichend dafür, dass der Summenscore

Die Suffizienz des Summenscores

$$x_{vo} = \sum_{i=1}^k x_{vi}$$

eine suffiziente Statistik für den Personenparameter θ_v darstellt, so dass die (bedingte) Wahrscheinlichkeit des Antwortmusters $x_v = (x_{v1}, \dots, x_{vk})$ bei gegebenem Score x_{vo} nur noch von den Itemparametern $\delta_1, \dots, \delta_k$ abhängt, vom Personenparameter θ_v jedoch unabhängig ist.

Was die Schätzung der Modellparameter betrifft, ermöglicht dies eine zweistufige Vorgehensweise, bei der zunächst die Itemparameter mittels der sog. bedingten (engl. *conditional*) Maximum-Likelihood-Methode (CML-Methode); Andersen, 1973) oder mittels der marginalen Maximum-Likelihood-Methode (MML-Methode; Thissen, 1982) geschätzt werden. Die Itemschwierigkeiten als bekannt voraussetzend, werden die Personenparameter anschließend mittels der gewichteten (engl. *weighted*) Maximum-Likelihood-Methode (WML-Methode; Warm, 1989) bestimmt.

Schätzung der Item- und Personenparameter

Conditional, Marginal und Weighted Maximum-Likelihood Methode

Weiterentwicklungen der Maximum-Likelihood Methode

Während bei der CML-Methode die bedingte Likelihood

$$CL(\mathbf{X}) = \prod_{v=1}^n \text{prob}\{(x_{v1}, \dots, x_{vk}) | x_{vo}\} \quad [9]$$

maximiert wird, die wegen der Suffizienz der x_{vo} von den Personenparametern unabhängig ist, behandelt die MML-Methode die Personenparameter als Realisationen einer Zufallsvariable Θ mit der Wahrscheinlichkeitsdichte $g(\theta)$ und maximiert die Likelihood

$$\begin{aligned} ML(\mathbf{X}) &= \prod_{v=1}^n \text{prob}\{(x_{v1}, \dots, x_{vk})\} \\ &= \prod_{v=1}^n \int_{-\infty}^{\infty} \text{prob}\{(x_{v1}, \dots, x_{vk}) | \theta\} g(\theta) d\theta \end{aligned} \quad [10]$$

Wegen der Suffizienz der Summenscores ist im Rasch-Modell aber

$$\text{prob}\{(x_{v1}, \dots, x_{vk})\} = \text{prob}\{(x_{v1}, \dots, x_{vk}) | x_{vo}\} \text{prob}(x_{vo}),$$

und daher

$$\begin{aligned} \text{ML}(X) &= \prod_{v=1}^n \text{prob}\{(x_{v1}, \dots, x_{vk}) | x_{vo}\} \text{prob}(x_{vo}) \\ &= \prod_{v=1}^n \text{prob}\{(x_{v1}, \dots, x_{vk}) | x_{vo}\} \prod_{v=1}^n \text{prob}(x_{vo}) \\ &= \text{CL}(X) \prod_{g=0}^k p_g^{n_g} \end{aligned}$$

worin $p_g = \text{prob}(x_{vo}=g)$ für die Wahrscheinlichkeit steht, mit der eine zufällig herausgegriffene V_p den Score g erzielt, und n_g die Anzahl der V_p n bezeichnet, deren Score gleich g ist. Da es sich bei den Scores um eine manifeste Variable handelt, kann die Wahrscheinlichkeit p_g unabhängig von den Item- und Personenparametern dargestellt werden, so dass der Logarithmus der marginalen Likelihood

$$\ln\{\text{ML}(X)\} = \ln\{\text{CL}(X)\} + \sum_{g=0}^k n_g \ln(p_g)$$

in zwei Summanden zerfällt, deren einer den Logarithmus der bedingten Likelihood darstellt und deren anderer von den Modellparametern unabhängig ist und daher beim Differenzieren wegfällt,¹ so dass

$$\frac{\partial \ln\{\text{ML}(X)\}}{\partial \delta_i} = \frac{\partial \ln\{\text{CL}(X)\}}{\partial \delta_i} \quad \text{für } i = 1, \dots, k.$$

Die MML-Methode führt im Falle des Rasch-Modells dann zu exakt denselben Schätzgleichungen für die Itemparameter wie die CML-Methode, deren asymptotische Eigenschaften von Andersen (1973) untersucht wurden: Normalverteilung, Erwartungstreue und Wirksamkeit.

Verzichtet man darauf, die marginale Likelihood in ihre Bestandteile zu zerlegen und berechnet sie unter Zusatzannahmen über die Wahrscheinlichkeitsdichte der Personenparameter nach For-

1. Zur Vorgehensweise bei der Konstruktion von Maximum-Likelihood-Schätzfunktionen siehe Kapitel 2.6.3.1.

mel [10], so sind die MML-Schätzer den CML-Schätzern immerhin noch *asymptotisch äquivalent*, wie aus einem Theorem von Pfanzagl (1993) hervorgeht.

Die zur Schätzung der Personenparameter angewandte WML-Methode beruht dagegen auf dem Bayes-Ansatz der Parameterschätzung, der im Unterschied zu den anderen Maximum-Likelihood-Verfahren *nicht* die Likelihood der Daten bei gegebenen Modellparametern maximiert, *sondern* die Likelihood der (Personen-)Parameter bei gegebenen Daten und Itemparametern.

In pragmatischer Hinsicht bedeutet die Suffizienz des Summenscores nichts anderes als die empirische Rechtfertigung der in der Psychodiagnostik üblichen Praxis, die Testleistung eines Probanden durch die Anzahl der gelösten Testaufgaben zu operationalisieren.

Pragmatische
Vorzüge des Rasch-
Modells

Indem der Summenscore die statistische Information über die vom Test gemessene latente Eigenschaft der Probanden enthält, kann er ohne Informationsverlust zur Beschreibung der Testleistung herangezogen werden – oder anders ausgedrückt: Wenn die Modellannahmen des Rasch-Modells erfüllt sind, kommt es ausschließlich darauf an, *wie viele* Aufgaben eine θ_v gelöst hat, nicht jedoch darauf, *welche* Aufgaben dies sind.

Da zwischen den Summenscores x_{v0} und den Personenparametern θ_v (lediglich) eine streng monoton wachsende Beziehung besteht, ist die Verwendung der Summenscores allerdings mit einer Einbuße bezüglich der Skalenqualität verbunden: Statt mit einer Differenzenskala, wie bei den Personenparametern, hat man es bei den Summenscores nur noch mit einer Ordinalskala der latenten Eigenschaft zu tun. Empfehlenswert ist es daher, den Gepflogenheiten der diagnostischen Praxis entgegenzukommen, indem man die Personenparameter mittels der linearen Transformationen

$$\Theta \rightarrow Z: \theta_v \rightarrow z_v = \frac{\theta_v - E(\Theta)}{\sigma(\Theta)}$$

und

$$Z \rightarrow IQ: z_v \rightarrow iq_v = 100 + 15z_v$$

in intervallskalierte z-Werte (mit Erwartungswert 0 und Standardabweichung 1) und IQ-Punkte (mit Erwartungswert 100 und Standardabweichung 15) umrechnet.

Der pragmatische Vorteil, den man damit erzielt, besteht darin, dass man es (im Unterschied zur traditionellen Berechnungsweise des Intelligenzquotienten) jetzt tatsächlich mit einer Intervallskala zu tun hat, deren Skalenwerte sich zudem (im Unterschied zu den Personenparametern) in derselben Größenordnung bewegen wie der IQ, für dessen Interpretation sich seit seiner Einführung durch William Stern (1920) gewisse Standards herausgebildet haben.

Spezifische
Objektivität

Neben ihren statistischen und pragmatischen Konsequenzen hat die Separierbarkeit der Parameter auch eine grundlegende wissenschaftstheoretische Bedeutung, wofür Georg Rasch (1968) den Terminus ‚spezifische Objektivität‘ geprägt hat. Das Rasch-Modell ist notwendig und hinreichend für die wechselseitige Unabhängigkeit der Vergleiche zwischen Messobjekten (Vpn) und Messinstrumenten (Items):

- Die Aussagen, welche über die Relationen zwischen $n = 1, 2, 3, \dots$ Vpn gewonnen werden können, sind davon unabhängig, welche Items ausgewählt¹ und dem Vergleich zugrundegelegt werden.
- Umgekehrt sind die Aussagen, welche über die Relation zwischen $k = 1, 2, 3, \dots$ Items gewonnen werden können, unabhängig davon, anhand welcher Personenstichprobe sie gewonnen werden.

Empirischer
Unterschiedsbegriff

So definiert z.B. der Vergleich zweier Vpn v und w , von denen genau eine das Item i gelöst hat, einen empirischen Unterschiedsbegriff, der von den Eigenschaften des jeweiligen Items unabhängig ist.

$$\ln \left(\frac{\text{prob}\{X_{vi} = 1 | x_{vi} + x_{wi} = 1\}}{\text{prob}\{X_{wi} = 1 | x_{vi} + x_{wi} = 1\}} \right) = \theta_v - \theta_w$$

Modellgeltungstest

Diese substantielle Eigenschaft des Rasch-Modells bildet die Grundlage, auf welcher die Modellgeltung überprüft werden kann. Das heißt: Auf welcher man prüfen kann, ob

1. Vorausgesetzt, die Items messen alle dieselbe latente Eigenschaft θ .

- die Items eines Tests alle dieselbe latente Dimension Θ messen,
- die Itemcharakteristiken die Form der in Gleichung [6] dargestellten logistischen Funktion haben, und
- die Testleistung eines Probanden daher durch die Anzahl der gelösten Aufgaben erschöpfend beschrieben werden kann.

Zu diesem Zweck kann man die Itemschwierigkeiten zunächst aus m verschiedenen Personenstichproben schätzen, z.B. aus einer Stichprobe von Minderbegabten, einer Stichprobe von Normalbegabten und einer Stichprobe von Hochbegabten. Sind die Modellvoraussetzungen des Rasch-Modells erfüllt und damit spezifisch objektive – d.h. von der Wahl der Personenstichprobe unabhängige – Messungen der Itemschwierigkeiten gewährleistet, dann müssen alle diese Parameterschätzungen $\delta_i^{(1)}$, $\delta_i^{(2)}$, ..., $\delta_i^{(m)}$ untereinander und mit der Schätzung der Itemparameter aus der Gesamtstichprobe $\delta_i^{(G)}$ übereinstimmen, so dass

$$\delta_i^{(1)} = \delta_i^{(2)} = \dots = \delta_i^{(m)} = \delta_i^{(G)}.$$

Trägt man die Parameterschätzungen in ein zweidimensionales Streudiagramm ein, dann müssen die Items im Falle der Modellgeltung alle auf einer Geraden liegen, die mit einem Anstieg von 45° durch den Ursprung des Koordinatensystems führt (vgl. Abb. 2.4.11).

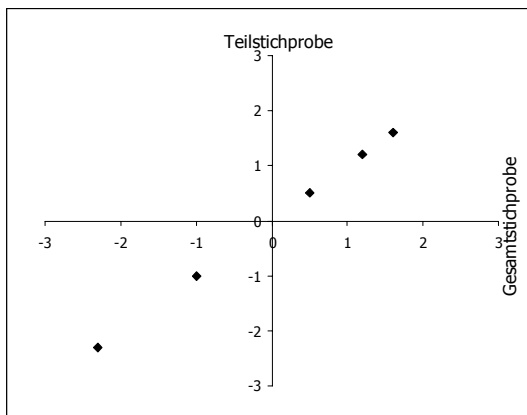
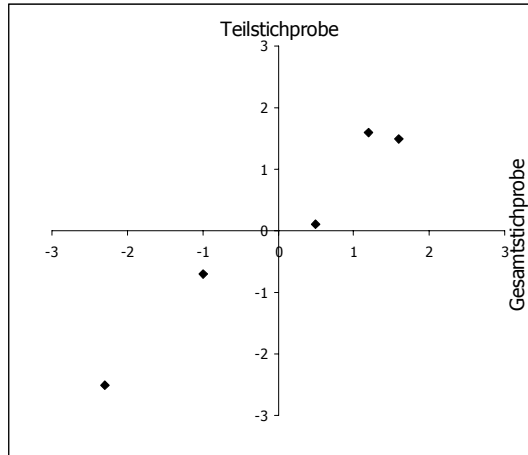


Abb. 2.4.11: Graphische Modellkontrolle des Rasch-Modells: Ideale Modellgeltung

In der Praxis werden wir – je nach Stichprobengröße – jedoch mehr oder minder starke Abweichungen von der 45° -Geraden finden (vgl. Abb. 2.4.12), so dass sich die Frage ergibt, ob es sich dabei lediglich um Zufallsfehler handelt oder ob die Abweichun-

gen statistisch signifikant sind und das Rasch-Modell verworfen werden muss.

Abb. 2.4.12: Graphische Modellkontrolle des Rasch-Modells: Geringfügige Zufallsfehler



Likelihood-Quotienten-Test

Um dies zu prüfen, kann man sich eines (bedingten) Likelihood-Quotienten-Tests (Andersen, 1973) bedienen, der

- die bedingte Likelihood der Datenmatrix unter Annahme der Modellgeltung (Nullhypothese: gleiche Itemparameter in allen Subpopulationen von Probanden)

$$L_0 = \text{CL}(\mathbf{X} | \delta_1^{(G)}, \dots, \delta_k^{(G)})$$

- und die bedingte Likelihood der Datenmatrix unter Annahme der Alternativhypothese (verschiedene Itemparameter in den verschiedenen Subpopulationen)

$$L_1 = \prod_{g=1}^m \text{CL}(\mathbf{X}^{(g)} | \delta_1^{(g)}, \dots, \delta_k^{(g)}) \quad \text{mit } \mathbf{X}^{(g)} = \text{Antwortmatrix der } g\text{-ten Teilstichprobe}$$

einander gegenüberstellt. Wegen $L_0 \leq L_1$ nimmt der Likelihood-Quotient

$$\lambda = \frac{L_0}{L_1}$$

stets Werte zwischen Null und Eins an, mit $\lambda = 1$ genau dann, wenn die Parameterschätzungen aus den Teilstichproben untereinander und mit den Parameterschätzungen aus der Gesamt-

stichprobe exakt übereinstimmen. Die asymptotisch nach χ^2 mit $df = (m-1)(k-1)$ Freiheitsgraden verteilte Testgröße

$$-2 \ln(\lambda) = 2 |\ln(L_0) - \ln(L_1)|$$

nimmt daher umso größere Werte an, je stärker die Modellannahmen des Rasch-Modells verletzt sind.¹ Man wird die Annahme der Modellgeltung daher verwerfen, wenn die Prüfgröße $-2 \ln(\lambda)$ auf einem Signifikanzniveau von $\alpha = 5\%$ den kritischen Wert der χ^2 -Verteilung mit $df = (m-1)(k-1)$ Freiheitsgraden übersteigt.

2.4.7 Exkurs: Informationsverarbeitungsmodelle der kognitiven Entwicklung

Während Latent-Trait- und Latent-Class-Modelle in der Psychometrie den State of the Art darstellen, konnten sie sich in der Psychodiagnostik bisher noch nicht entsprechend durchsetzen. Dies hat u.a. damit zu tun, dass

Psychometrie und Psychodiagnostik

1. viele der gängigen Testverfahren den messtheoretischen Voraussetzungen nicht gerecht werden, welche erforderlich sind, um den Schluss von manifesten Itemantworten auf eine quantitative latente Dimension zu rechtfertigen; während andererseits
2. die Annahme einer qualitativen latenten Variable nicht nur in theoretischer Hinsicht ein Abrücken von quantitativ gedachten Eigenschaftsbegriffen wie Intelligenz etc. erfordern würde, sondern auch die Praxis des Psychodiagnostikers verändern müsste, der es seit einem Jahrhundert gewohnt ist, das Testergebnis seiner Probanden durch die Anzahl der gelösten Aufgaben (Summenscore) darzustellen.

Bisherige Erfahrungen mit der Anwendung von Latent-Trait-Modellen in der Testanalyse und Testkonstruktion lassen jedenfalls den Schluss zu, dass es so umfassende latente Eigenschaftsdimensionen wie z.B. Intelligenz, die etwa William Stern (1920, 2) als die "allgemeine geistige Anpassungsfähigkeit an neue Aufgaben und Bedingungen des Lebens" umschrieben hatte, in der empirischen Wirklichkeit nicht gibt.

Gibt es eine generelle Intelligenz?

1. Vgl. dazu und zu weiteren Modellgeltungstests z.B. das Lehrbuch von Rost (1996). Dort finden sich auch genauere Details zu den oben skizzierten Schätzverfahren sowie zu verschiedenen Varianten und Weiterentwicklungen des Rasch-Modells. Zu den Grundbegriffen und der Logik des statistischen Hypothesentests vgl. Kapitel 2.6.5.

Was sich mithilfe von Tests messen lässt, sind (allenfalls) sehr spezielle Fähigkeiten zur Lösung sehr spezifischer Aufgabenarten, die inhaltlich eher noch homogener sind als die ohnedies schon sehr homogenen Subtests der gebräuchlichen Intelligenztests.

Andererseits korrelieren solche Subtests alle untereinander und auch mit Alltagsleistungen wie z.B. dem Schulerfolg, was bereits Charles Spearman (1904) zu der Überlegung veranlasst hatte, dass allen diesen teils komplexen, teils einfacheren Leistungen eine gemeinsame Fähigkeit zugrunde liegen müsse: die generelle Intelligenz. Im Gegensatz dazu vertrat Luis Thurstone (siehe Thurstone & Thurstone, 1941) die Auffassung, dass es sieben unabhängige Primärfähigkeiten seien, welche die Intelligenzleistungen ausmachen (vgl. Tab. 2.4.4).

Tab. 2.4.4: Thurstones sieben Primärfähigkeiten und ihre Operationalisierung

Primärfähigkeiten	Operationalisierungen
1. Verbales Verständnis	Wortschatz, verbale Analogien, Textverständnis, Rechtschreibung
2. Wortflüssigkeit	Schnelles Produzieren von Wörtern, Lösen von Anagrammen, Bilden von Reimen oder Finden von Benennungen
3. Zahlen	Geschwindigkeit und Genauigkeit bei einfachen Rechenaufgaben
4. Raum	Räumliches Vorstellungsvermögen, Orientierungsleistungen, Verständnis mechanischer Bewegungen und technischer Instrumente
5. Gedächtnis	Behalten gelernter Assoziationen zwischen Paaren von Wörtern, Zahlen, Bildern oder Figuren
6. Wahrnehmungsgeschwindigkeit	Geschwindigkeit bei der Identifikation visueller Konfigurationen, Erkennen von Gleichheiten und Ungleichheiten
7. Induktives Denken	Schlussfolgerndes Denken, Auffinden allgemeiner Regeln, z.B. bei Reihenfortsetzungsaufgaben

Sowohl Spearman als auch Thurstone meinten, ihre Theorie empirisch bestätigt zu finden. Die Methoden der Datenanalyse waren jedoch z.T. fragwürdig, und auch nach weiteren sechs Jahrzehnten empirischer Intelligenzforschung ist die Kontroverse zwischen Spearman und Thurstone immer noch unentschieden, was insbesondere der Stichprobenabhängigkeit der faktorenanalytischen Methodik geschuldet ist, mittels derer man die Frage zu

entscheiden versucht hat. Rost (1997, 545) kommt daher zu dem Schluss, dass die Frage nach der Anzahl der Intelligenzfaktoren auch prinzipiell unentscheidbar zu sein scheint.

Die Faktorenanalyse

Die Faktorenanalyse ist ein statistisches Verfahren, mit dem man die Matrix der Interkorrelationen von m manifesten Variablen durch eine geringere Anzahl von $q < m$ linear unabhängigen (d.h. miteinander nicht korrelierenden) latenten Variablen, den sogenannten Faktoren, zu beschreiben versucht, indem man die manifesten Variablen (X_{o1}, \dots, X_{om}) als Linearkombination der latenten Variablen (Y_{o1}, \dots, Y_{oq}) darstellt:

$$X_{ot} = a_{t1}Y_{o1} + a_{t2}Y_{o2} + \dots + a_{tq}Y_{oq}.$$

Je homogener die Personenstichprobe ist, desto niedriger fallen jedoch die Korrelationen zwischen den verschiedenen Tests aus und desto mehr unabhängige Faktoren werden benötigt, um die Korrelationsmatrix zu beschreiben. Je heterogener die Personenstichprobe ist, desto höher sind die Interkorrelationen und desto weniger Faktoren werden benötigt.

Für die Erklärung der Intelligenzleistungen ist es ohnedies unerheblich, durch wie viele Faktoren sich die Intelligenz am besten beschreiben lässt, und gelegentlich drängt sich geradezu die Frage auf, ob es bei dem ganzen Streit nicht letztlich nur darum geht, wer denn nun die "richtige" Vermögenspsychologie vertritt. Wollen wir der vermögenspsychologischen Falle entgehen, kommen wir nicht umhin, einzugestehen, dass es sich bei der Intelligenz lediglich um einen Sammelnamen für eine Vielzahl von untereinander korrelierenden Einzelleistungen handelt, nicht jedoch um eine quantitative Dimension, auf der man die Probanden anordnen kann und/oder durch welche diese Leistungen erklärt werden können.

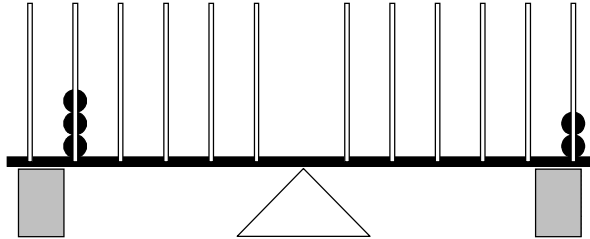
Darüber hinaus stellt sich schon bei relativ spezifischen kognitiven Leistungen die Frage, ob es denn überhaupt zweckmäßig ist, diese auf einer quantitativen Dimension messen zu wollen. Um dies zu verdeutlichen, betrachten wir im Folgenden ein Experiment von Siegler (1976) zur Entwicklung des formal-operationalen Denkens bei Kindern. Mit dem Experiment soll gezeigt werden, dass die der Annahme einer Stufenentwicklung der Intelligenz (Pi-

Stichproben-
abhängigkeit
faktorenanalytischer
Ergebnisse

Stufenentwicklung
der Intelligenz

aget) zugrundeliegenden Leistungssprünge durch einen sukzessiven Wissenszuwachs erklärt werden können. Um dies nachzuweisen, verwendet Siegler einen Typ von Problemlösungsaufgaben, der schon von Inhelder & Piaget (1958) verwendet wurde: die Balkenwaage (vgl. Abb. 2.4.13).

Abb. 2.4.13: Die Balkenwaage



"Welche Seite neigt sich nach unten, wenn man die Holzklötzchen entfernt?"

Schrittweise
Akkumulation von
Wissen

Dabei unterscheidet Siegler sechs verschiedene Aufgabentypen (vgl. Tab. 2.4.5) und beschreibt die schrittweise Akkumulation von Wissen über die physikalischen Gesetzmäßigkeiten, welche das Verhalten der Balkenwaage bestimmen, in vier Entwicklungsstufen, denen jeweils eine bestimmte Lösungsstrategie entspricht, die in Form eines Flussdiagramms (vgl. Abb. 2.4.14-2.4.17) dargestellt werden kann.

Abb. 2.4.14: Die erste Entwicklungsstufe

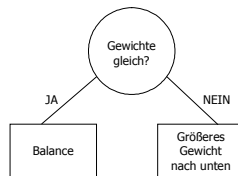
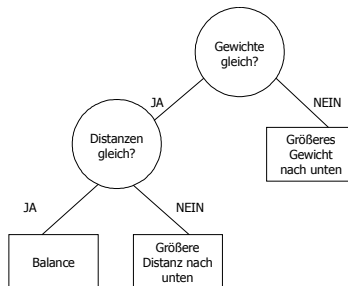


Abb. 2.4.15: Die zweite Entwicklungsstufe



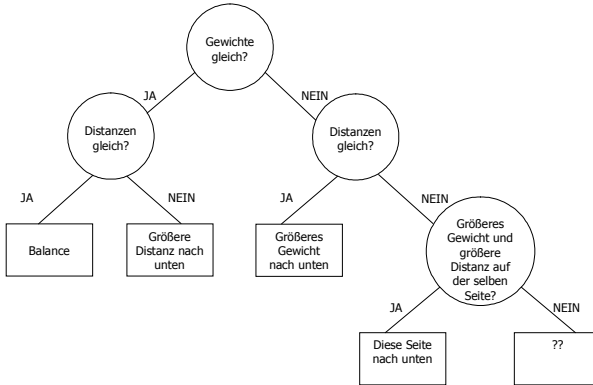


Abb. 2.4.16: Die dritte Entwicklungsstufe

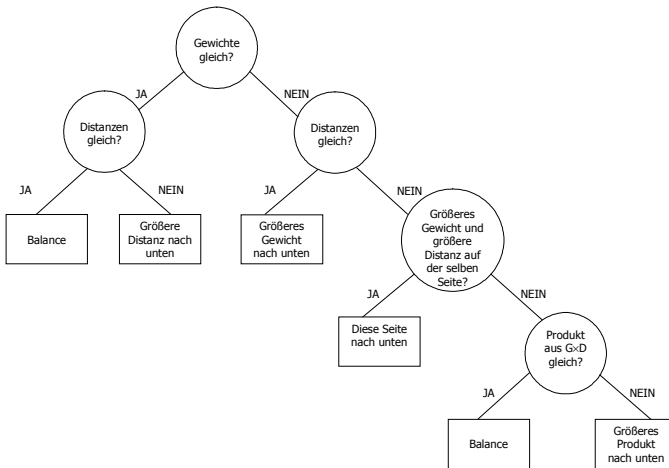


Abb. 2.4.17: Die vierte Entwicklungsstufe

Auf der ersten Entwicklungsstufe wissen die Kinder lediglich, dass das Verhalten der Balkenwaage etwas mit Gewichten zu tun hat. Das größere Gewicht entscheidet.

Die vier Entwicklungsstufen

Auf der zweiten Entwicklungsstufe entscheidet ebenfalls das größere Gewicht. Bei gleichen Gewichten wird aber auch bereits die Distanz beachtet.

Auf der dritten Entwicklungsstufe spielen Distanz und Gewicht eine gleichwertige Rolle. Widersprechen sie einander, kann das Kind nur raten.

Auf der vierten Entwicklungsstufe spielen Distanz und Gewicht eine gleichwertige Rolle. Widersprechen sie einander, so ent-

scheidet das größere Drehmoment (= Produkt aus Gewicht \times Distanz).

Prognose der Antwortmuster und Diagnose der Entwicklungsstufen

Auf Grundlage der Lösungsstrategien können für die verschiedenen Entwicklungsstufen charakteristische Antwortmuster prognostiziert (vgl. Tab. 2.4.5) und die jeweiligen Entwicklungsstufen der Kinder diagnostiziert werden, indem man ihre tatsächlichen Antwortmuster mit den prognostizierten Mustern vergleicht.

Tab. 2.4.5: Prognostizierte Antwortmuster für sechs Aufgabentypen und vier Entwicklungsstufen. Richtige Antworten sind grau unterlegt¹

Aufgabentyp	Beispiel				Stufe			
	Links		Rechts					
	G	D	G	D	1	2	3	4
Balance (B)	2	3	2	3	B	B	B	B
Gewicht (G)	1	1	5	1	G	G	G	G
Distanz (D)	3	2	3	3	B	D	D	D
Konflikt-Gewicht (K-G)	1	3	4	1	G	G	?	G
Konflikt-Distanz (K-D)	3	3	6	1	G	G	?	D
Konflikt-Balance (K-B)	2	2	4	1	G	G	?	B

Wie man unmittelbar sieht, handelt es sich dabei um eine äußerst präzise Form der psychologischen Diagnostik, die sich nicht damit begnügt, die Kinder als mehr oder weniger leistungsfähig zu beurteilen, sondern die zugleich das Zustandekommen der Testleistung erklärt – und vor allem genaue Anhaltspunkte dafür gibt, was ein Kind bereits weiß, was es noch nicht weiß und worauf man didaktisch aufbauen kann, wenn man die Leistungsfähigkeit eines Kindes durch Training verbessern will.

Durch eine quantitative Diagnostik, welche die Testleistung durch den Summenscore beschreibt, hätte dies nicht geleistet werden können. Ja man hätte noch nicht einmal die Kinder der zweiten Stufe von denen der dritten unterscheiden können, wie man sofort sieht, wenn man aus Tab. 2.4.5 die Lösungswahrscheinlichkeiten pro Stufe \times Aufgabentyp ableitet und daraus für jede Entwicklungsstufe (g) die erwartete Anzahl richtiger Antworten $E(X_{g0})$ berechnet (vgl. Tab. 2.4.6). Die Kinder der Stufen 2 und 3

1. Balance-Aufgaben: Gewichte gleich und Distanzen gleich. Gewicht-Aufgaben: Distanzen gleich, Gewichte verschieden. Distanz-Aufgaben: Gewichte gleich, Distanzen verschieden. Konflikt-Aufgaben: Gewichte und Distanzen gegenläufig. Konflikt-Gewicht: höheres Gewicht überwiegt größere Distanz. Konflikt-Distanz: größere Distanz überwiegt höheres Gewicht. Konflikt-Balance: höheres Gewicht und größere Distanz kompensieren sich gegenseitig.

unterscheiden sich ausschließlich dadurch, welche Aufgaben sie lösen, nicht jedoch durch die Anzahl der gelösten Aufgaben.

Stufe	Aufgabentyp						$E(X_{go})$
	B	G	D	K-G	K-D	K-B	
1	1.00	1.00	0.00	1.00	0.00	0.00	3
2	1.00	1.00	1.00	1.00	0.00	0.00	4
3	1.00	1.00	1.00	0.33	0.33	0.33	4
4	1.00	1.00	1.00	1.00	1.00	1.00	6

Tab. 2.4.6: Stufenspezifische Lösungswahrscheinlichkeiten und erwartete Anzahl richtiger Antworten (unter Zugrundelegung einer Ratewahrscheinlichkeit von $p=0.33$). Unterschiede in der Testleistung von Kindern der Stufen 2 und 3 sind grau unterlegt

2.4.8 Latent-Class-Modelle¹

Dass man in der Einstellungsdiagnostik nicht unbedingt davon ausgehen kann, dass sich die Probanden auf einer latenten Einstellungsdimension anordnen lassen, haben wir bereits in Kapitel 2.4.1 diskutiert. Auch in der klinischen Psychologie geht man traditionell davon aus, dass es nicht die Anzahl der Krankheits-symptome ist, welche ein spezifisches Krankheitsbild ausmacht, sondern ein typisches Symptommuster (Syndrom). Das zuletzt diskutierte Experiment von Siegler (1976) schließlich hat gezeigt, dass die Antwortmuster selbst in der Diagnostik kognitiver Prozesse ggf. informativer sein können als die Leistungsmenge (Summenscore).

Qualitative latente Variablen

Allen diesen Beispielen ist gemeinsam, dass von manifesten Variablen (den Antworten der Probanden auf die Items eines Fragebogens, den vom behandelnden Arzt festgestellten Krankheits-symptomen und/oder den Ergebnissen von Problemlöseprozessen) auf eine qualitative latente Variable geschlossen wird, welche die Probanden in verschiedene Klassen (qualitativ verschiedene rassistische Einstellungen, klinische Syndrome und/oder Stufen der kognitiven Entwicklung) einteilt.

Am Beispiel des Siegler-Experimentes haben wir auch schon gesehen, wie dieser Schluss idealiter erfolgen kann, wenn die Klassen als solche bereits bekannt sind. Ähnlich wie bei der Guttman-Skala handelt es sich bei Sieglers Modell jedoch (abgesehen von den Ratewahrscheinlichkeiten auf Stufe 3) um ein deterministi-

1. Dieser Abschnitt kann ohne Verlust der Kontinuität übersprungen werden.

sches Modell, das sich nur dank einer äußerst sorgfältigen und empathischen Experimentiertechnik praktisch bewähren konnte.

Dies wirft die Frage nach der Identifizierbarkeit von latenten Klassen auf, wenn (1) die Daten von Zufallsschwankungen überlagert sind, (2) die Klassenzugehörigkeit der Probanden nicht bekannt ist und wenn wir (3) auch noch nicht wissen, wie die latenten Klassen denn eigentlich aussehen.

Dies zu untersuchen, ist Gegenstand der auf Lazarsfeld (1950) zurückgehenden Latent-Class-Analyse (LCA).

Das Modell der Latent-Class-Analyse

Allgemein formuliert, haben wir es in diesem Fall mit k manifesten Variablen ($i = 1, \dots, k$) und h latenten Klassen ($g = 1, \dots, h$) zu tun:

- Die manifesten Variablen sind qualitativ und können jeweils m_i verschiedene Werte annehmen, die wir o.B.d.A. mit $x = 0, 1, \dots, m_i - 1$ bezeichnen.
- Das Charakteristikum, wodurch sich die latenten Klassen voneinander unterscheiden, besteht in den klassenspezifischen Kategorienwahrscheinlichkeiten $p_{gix} = \text{prob}\{X_{vi} = x | v \in g\}$, welche die Wahrscheinlichkeit beschreiben, mit der eine V_p aus Klasse g in Variable i den Wert x erzielt.
- Die Wahrscheinlichkeit $\text{prob}\{v \in g\}$, mit der eine zufällig herausgegriffene V_p der Klasse g angehört, entspricht der (relativen) Klassengröße p_g .

Für die marginale Likelihood der Datenmatrix folgt daraus die Modellgleichung der Latent-Class-Analyse:

$$\begin{aligned} \text{ML}(\mathbf{X}) &= \prod_{v=1}^n \left(\sum_{g=1}^h \text{prob}\{v \in g\} \text{prob}\{(x_{v1}, \dots, x_{vk}) | v \in g\} \right) & [11] \\ &= \prod_{v=1}^n \left(\sum_{g=1}^h \text{prob}\{v \in g\} \prod_{i=1}^k \text{prob}\{X_{vi} = x_{vi} | v \in g\} \right) \end{aligned}$$

Ein-Klassen-Lösung und saturiertes Modell

Weder die Klassengrößen p_g noch die klassenspezifischen Kategorienwahrscheinlichkeiten p_{gix} sind bekannt, und auch von der Anzahl der Klassen, die zur Beschreibung der Daten erforderlich sind, wissen wir lediglich, dass sie zwischen $1 \leq h \leq h_{\max}$ liegt, worin

$$h_{\max} = \prod_{i=1}^k m_i$$

die Anzahl der möglichen Antwortmuster (x_{v1}, \dots, x_{vk}) darstellt.

Die gesuchte Lösung der LCA liegt also irgendwo zwischen zwei Extremen, wonach entweder

- alle Daten aus derselben homogenen Verteilung stammen (Ein-Klassen-Lösung; $h=1$), wie dies bei der traditionellen statistischen Analyse angenommen wird, oder
- jedes mögliche Antwortmuster eine eigene Klasse darstellt (saturiertes Modell; $h=h_{\max}$).

In beiden Extremfällen ist die Schätzung der Modellparameter relativ einfach zu bewerkstelligen:

- Bei der Ein-Klassen-Lösung ist die Klassengröße per definitionem gleich 100%, und die Kategorienwahrscheinlichkeiten können durch Auszählen der entsprechenden relativen Häufigkeiten geschätzt werden.
- Beim saturierten Modell ist jedes Muster $(x_{v1}, \dots, x_{vk}) = \mathbf{x}_v$ eine eigene Klasse, so dass die Wahrscheinlichkeiten

$$\text{prob}\{(x_{v1}, \dots, x_{vk}) | v \in g\} = \begin{cases} 1 & \text{wenn } (x_{v1}, \dots, x_{vk}) \equiv g \\ 0 & \text{sonst} \end{cases}$$

In jeder Klasse kommt somit das entsprechende Muster mit Sicherheit vor, und die Klassengrößen $\text{prob}\{v \in g\}$ können durch die relativen Musterhäufigkeiten geschätzt werden.

Auf ähnlich einfache Weise könnten wir die Modellparameter p_g und p_{gix} auch für $1 < h < h_{\max}$ Klassen schätzen, wenn es sich um manifeste Klassen handelte, d.h. wenn a priori bekannt wäre, welche V_{pn} jeweils derselben Klasse angehören. Dann gilt:

$$p_g \approx n_g/n$$

und

$$p_{gix} \approx n_{gix}/n_g,$$

worin n_g und n_{gix} die entsprechenden absoluten Häufigkeiten bezeichnen. Gerade dies ist bei der Latent-Class-Analyse jedoch nicht der Fall. Wir wissen nicht a priori, welche V_{pn} zur selben Klasse gehören, und die Häufigkeiten n_g und n_{gix} können daher nicht einfach ausgezählt werden. Sie sind unbekannt.

Eine zufriedenstellende Lösung des Problems wurde erst von Goodman (1974, 1979) gefunden. Später hat Andersen (1982) gezeigt, dass es sich dabei um einen Spezialfall des EM-Algorithmus handelt, der in seiner allgemeinen Form von Dempster et al. (1977) untersucht wurde.

Die Schätzung der Modellparameter für $1 < h < h_{\max}$ Klassen

Der EM-Algorithmus

Der EM-Algorithmus

Der EM-Algorithmus ist ein iteratives Verfahren zur ML-Schätzung der Modellparameter.

- Ausgehend von vorläufigen Schätzwerten (Startwerten) der Parameter p_g und p_{gix} werden zunächst im sogenannten E-Schritt (Erwartungswerte) die bei diesen Parameterwerten erwarteten Häufigkeiten e_g und e_{gix} ermittelt.
- Aus diesen werden im darauffolgenden M-Schritt (Maximum Likelihood) verbesserte Schätzwerte der unbekanntnen Parameter $p_g = e_g/n$ und $p_{gix} = e_{gix}/e_g$ berechnet.

Die verbesserten Schätzwerte werden als neue vorläufige Schätzwerte in den E-Schritt eingesetzt und das ganze Verfahren so lange wiederholt, bis die (marginale) Likelihood [11] nicht mehr weiter wächst.

Auf den ersten Blick scheint es sich dabei um ein zirkuläres Vorgehen zu handeln. Tatsächlich ist dies jedoch nicht der Fall, da die erwarteten Häufigkeiten nicht einfach über die Gleichungen $e_g = np_g$ und $e_{gix} = e_g p_{gix}$ berechnet werden, sondern über die Membership-Wahrscheinlichkeiten

$$\begin{aligned} p_{g,xv} &= \text{prob}\{v \in g | (x_{v1}, \dots, x_{vk})\} & [12] \\ &= \frac{\text{prob}\{(v \in g) \wedge (x_{v1}, \dots, x_{vk})\}}{\text{prob}\{(x_{v1}, \dots, x_{vk})\}} \\ &= \frac{\text{prob}\{v \in g\} \text{prob}\{(x_{v1}, \dots, x_{vk}) | v \in g\}}{\sum_{d=1}^h \text{prob}\{v \in d\} \text{prob}\{(x_{v1}, \dots, x_{vk}) | v \in d\}} \end{aligned}$$

so dass bei der Ermittlung der erwarteten Häufigkeiten

$$e_g = \sum_{v=1}^n p_{g,xv}$$

und

$$e_{gix} = \sum_{v: x_{vi}=x} p_{g,xv}$$

jedes Mal erneut Information aus den Daten aufgenommen wird. Diese Vorgehensweise ist dem hermeneutischen Zirkel in der interpretativen Sozialforschung nicht unähnlich.

Mit der Lösung des Schätzproblems ist aber noch nicht die Frage beantwortet, *wie viele* latente Klassen zu unterscheiden sind. In der Praxis geht man so vor, dass man die LCA zunächst probe-weise für die Klassenanzahlen $h = 1, 2, 3, \dots$ berechnet und daraus jenes LCA-Modell (d.h. jene Klassenanzahl) auswählt, welches die Daten am besten beschreibt.

Die Bestimmung der Klassenanzahl

Man kann aber nicht einfach jene Klassenzahl auswählen, unter welcher die Datenmatrix die größte (marginale) Likelihood besitzt. Da die Likelihood eine monoton wachsende Funktion der Anzahl der latenten Klassen ist, würde dies stets zu einer Entscheidung zugunsten des saturierten Modells führen.

Je mehr latente Klassen man annimmt, desto mehr Modellparameter sind aus den Daten zu schätzen; je mehr Modellparameter zu schätzen sind, desto weniger wird die Verteilung der Datenmatrix restringiert; und je weniger die Verteilung der Datenmatrix restringiert wird, desto größer wird ihre Likelihood. Als Kriterium für die Güte der Beschreibung müssen wir daher ein Maß wählen, das die Likelihood der Datenmatrix und die Anzahl der zu schätzenden Modellparameter $n(P_h)$ zueinander in Beziehung setzt. Ein solches Maß ist z.B. der AIC-Index (Akaike, 1987):

$$\text{AIC} = 2 |\ln\{\text{ML}(\mathbf{X})\}| + 2 n(P_h),$$

worin sich die Anzahl der (unabhängigen) Modellparameter nach

$$n(P_h) = h - 1 + h \sum_{i=1}^k (m_i - 1)$$

berechnet. Darin ist $h-1$ die Anzahl der unabhängigen Klassen-Größen und $h \sum_{i=1}^k (m_i - 1)$ die Anzahl der unabhängigen Kategorien-wahrscheinlichkeiten.¹

Da die Likelihood $0 \leq \text{ML}(\mathbf{X}) \leq 1$ eine Zahl zwischen Null und Eins ist, ist ihr natürlicher Logarithmus $\ln\{\text{ML}(\mathbf{X})\} \leq 0$ und daher dem Betrage nach umso kleiner, je größer die Likelihood ist. Folglich ist der AIC-Index umso kleiner,

- je genauer die Daten beschrieben werden (große Likelihood) und
- je sparsamer die Beschreibung ist (geringe Parameterzahl).

1. Entsprechend ist im Falle der Ein-Klassen Lösung $n(P_1) = \sum m_i - k$ und im Falle des saturierten Modells ist $n(P_{\text{sat}}) = h_{\text{max}} - 1$.

Man wird daher jenes LCA-Modell auswählen, welches den *kleinsten* AIC-Index hat.

Vergleiche des so identifizierten "besten" LCA-Modells (a) mit der Ein-Klassen-Lösung und (b) mit dem saturierten Modell können mittels Likelihood-Quotienten-Tests durchgeführt werden und geben Aufschluss darüber, ob (a) die Verteilung der Daten überhaupt eine Mischverteilung mehrerer latenter Klassen darstellt bzw. (b) ob das ausgewählte LCA-Modell die Daten signifikant schlechter beschreibt als das saturierte Modell, welches die bestmögliche Beschreibung der Daten leistet.

Allgemein kann der Vergleich zweier LCA-Modelle mit $h = y$ bzw. $h = z$ latenten Klassen ($y < z$) mittels der asymptotisch nach χ^2 mit $df = n(P_z) - n(P_y)$ Freiheitsgraden verteilten Testgröße

$$-2 \ln(\lambda) = 2 |\ln(ML_y) - \ln(ML_z)|$$

erfolgen, worin ML_y die (marginale) Likelihood unter Annahme der Nullhypothese ($h = y$) bezeichnet. ML_z bezeichnet die Likelihood unter Annahme der Alternativhypothese ($h = z$).

Voraussetzung für die Anwendung des Likelihood-Quotienten-Tests ist es allerdings, dass das unter der Nullhypothese postulierte Modell einen echten Spezialfall des Modells unter der Alternativhypothese darstellt, weshalb praktisch nur der Vergleich mit der Ein-Klassen Lösung und/oder der Vergleich mit dem saturierten Modell in Frage kommt.

Individualdiagnostik

Hat man erst einmal das LCA-Modell identifiziert, welches die Daten am besten beschreibt, kann die Zuordnung der V_p zu den latenten Klassen auf Grundlage ihrer Membership-Wahrscheinlichkeiten [12] erfolgen, so dass jede V_p der latenten Klasse zugeordnet wird, der sie aufgrund ihres Antwortmusters am wahrscheinlichsten angehört. Praktische Erfahrungen mit der Latent-Class-Analyse haben gezeigt, dass dabei in den meisten Fällen sehr klare diagnostische Entscheidungen mit großer Zuordnungssicherheit getroffen werden können.

Kontrollfragen

- Was versteht man unter einer Messung, welcher Art ist dabei die Zuordnung von Maßzahlen zu Messobjekten, und welchen Sinn hat das Ganze?
- Worin besteht der Unterschied zwischen empirischen Relationen und

- numerischen Relationen?
- Was versteht man unter der Homomorphie der Abbildung empirischer Relationen auf numerische Relationen?
 - Nennen Sie einige Beispiele für qualitative Variablen!
 - Auf welcher Art von Skala können qualitative Variablen gemessen werden?
 - Worin besteht der Unterschied zwischen dichotomen und polytomen Variablen?
 - Was versteht man unter einer binären Variable?
 - Worin besteht der Unterschied zwischen qualitativen und quantitativen Variablen?
 - Welche Mindestvoraussetzung muss gegeben sein, damit wir es mit einer quantitativen Variable zu tun haben?
 - Erläutern Sie anhand eines Beispiels, dass bloße Quantifizierung von Variablen noch nicht deren Messung bedeutet!
 - Erläutern Sie anhand des Beispiels einer Rassismus-Skala, wie die empirischen Ordnungsrelationen $>^{\circ}$ definiert sind, auf denen die Guttman-Skala beruht!
 - Welche Voraussetzung müssen diese Ordnungsrelationen erfüllen, damit die Probanden auf einer quantitativen Dimension (Rassismus) angeordnet werden können?
 - Welche Konsequenz hat dies für die zulässigen Antwortmuster der Probanden?
 - Wie lässt sich diese Gesetzmäßigkeit als Funktion der Rangplätze darstellen, welche die Probanden und die Items auf der gemeinsamen (Rassismus-)Skala einnehmen?
 - Wie errechnen sich die Rangplätze der Probanden, wenn man die nichtnegativen ganzen Zahlen als Skalenwerte verwendet?
 - Wie errechnen sich in diesem Fall die (Schwierigkeits-)Rangplätze der Items?
 - Nennen Sie die 6 wichtigsten Skalentypen und geben Sie Beispiele dafür!
 - Geben Sie für jeden dieser Skalentypen an, welche Relationen durch Skalen dieses Typs abgebildet werden und erläutern Sie dies durch Beispiele!
 - Was wird von der allgemeinen Messtheorie unter dem Stichwort ‚Repräsentationsproblem‘ untersucht? Geben Sie ein Beispiel!
 - Womit befasst sich die angewandte Messtheorie unter dem Stichwort ‚Skalierungsproblem‘?
 - Was untersucht die allgemeine Messtheorie unter dem Stichwort ‚Eindeutigkeitsproblem‘?
 - Wodurch ist der Skalentyp einer Messung definiert?
 - Welche Maßzahltransformationen sind bei
 - a) Nominalskalen,
 - b) Ordinalskalen,
 - c) Intervallskalen,
 - d) Differenzskalen,
 - e) Rationalskalen und
 - f) Absolutskalen
 zulässig?

- Stellen Sie die Hierarchie der o.g. Skalentypen dar und erläutern Sie die Konsequenzen, welche sich daraus ergeben, wenn auf eine Skala eine Maßzahltransformation angewendet wird, die nur für einen niedrigeren Skalentyp zulässig ist!
- Worin besteht das Bedeutsamkeitsproblem?
- Geben Sie für jeden der o.g. Skalentypen an,
 - a) welche Art von Aussagen bedeutsam sind, und
 - b) welche statistischen Verfahren zulässig sind!
- Warum stellt die routinemäßige Anwendung parametrischer statistischer Verfahren eine Fehlerquelle empirischer psychologischer Untersuchungen dar?
- Worin besteht der Unterschied zwischen manifesten und latenten Variablen?
- Geben Sie ein Beispiel dafür, dass manifeste Variablen einen anderen Skalentyp haben (können) als die latente Variable, welche durch sie operationalisiert wird! Welche der beiden hat den höheren Skalentyp?
- Welche Art von statistischen Verfahren ist für die statistische Weiterverarbeitung von ordinalen Daten angemessen?
- Was versteht man unter der Invarianz der Normalverteilung gegenüber positiven linearen Transformationen?
- Welche Beziehung besteht zwischen den Verteilungseigenschaften und dem Skalentyp normalverteilter Variablen?
- Was muss gegeben sein, damit man von manifesten Indikatoren auf latente Variablen schließen kann? Geben Sie ein Beispiel!
- Welchen Zusammenhang postuliert die differentielle Psychologie zwischen der Intelligenz der Probanden und ihrer Testleistung in einem Intelligenztest?
- Erläutern Sie am Beispiel der Intelligenzmessung den Terminus ‚Itemcharakteristik‘: Was beschreibt die Itemcharakteristik, welche Form hat sie, und welche Ordnungsrelation bezüglich der Intelligenz der Vpn wird dadurch definiert?
- Mit welchen Arten von quantitativen und qualitativen, latenten und manifesten Variablen haben wir es dabei zu tun, und was kann über ihre Skaleneigenschaften ausgesagt werden?
- Auf welcher Art von Skala werden die Itemschwierigkeiten und die Intelligenz der Vpn gemessen, wenn die Itemcharakteristiken die Form der logistischen Funktion haben, und warum?
- Welcher Zusammenhang besteht zwischen dem Skalentyp der manifesten Variablen und dem Skalentyp der latenten Variablen, für welche sie als Indikatoren dienen?
- Welche terminologischen Regeln bestehen zwischen den Termini ‚Latent-Trait-Modell‘, ‚Latent-Class-Modell‘, ‚Item-Response-Modell‘, ‚klassische Testtheorie‘, ‚stochastische Testmodelle‘ und ‚probabilistische Testmodelle‘?
- Was versteht man unter Latent-Trait-Modellen?
- Worin besteht der Unterschied zwischen Latent-Trait- und Latent-Class-Modellen?
- Worin besteht der Unterschied zwischen Item-Response-Modellen und der klassischen Testtheorie?
- Welche beiden Arten von stochastischen Testmodellen lassen sich unterscheiden?

- Welche (zweistellige) empirische Ordnungsrelation wird im Rasch-Modell auf die numerische Relation $\theta_v > \theta_w$ abgebildet?
- Worin besteht im Kontext von Latent-Trait-Modellen das Skalierungsproblem?
- Welche statistischen Schätzprobleme ergeben sich daraus, und warum?
- Was versteht man im Rasch-Modell unter der Separierbarkeit der Modellparameter? Auf welcher Eigenschaft der logistischen Funktion beruht sie?
- Welcher zweistufigen Vorgehensweise bedient man sich bei der Schätzung der Modellparameter im Rasch-Modell?
- Worin besteht der Unterschied zwischen der CML- und der MML-Methode? Weshalb führen die beiden Verfahren im Falle des Rasch-Modells zu exakt denselben Schätzgleichungen?
- Welche asymptotischen Eigenschaften haben CML-Schätzungen?
- Worin besteht der Bayes-Ansatz der Parameterschätzung, auf dem die WML-Methode beruht?
- Welche pragmatischen Vorzüge bietet die im Rasch-Modell gegebene Suffizienz der Summenscores?
- Welche Art von Beziehung besteht im Rasch-Modell zwischen den Summenscores und den Personenparametern? Welche Skaleneigenschaft haben die Summenscores? Welche Skaleneigenschaft haben die Personenparameter?
- Wie kann man die Personenparameter in IQ-Punkte umrechnen? Welche Skaleneigenschaft haben die so errechneten IQ-Punkte? Welche Verteilungseigenschaften haben sie?
- Was versteht man unter der im Rasch-Modell gewährleisteten spezifischen Objektivität?
- Welcher empirische Unterschiedsbegriff wird im Rasch-Modell auf die Differenz $\theta_v - \theta_w$ abgebildet?
- Nennen Sie einige der Annahmen, die durch den Modellgeltungstest des Rasch-Modells geprüft werden!
- Auf welcher Eigenschaft des Rasch-Modells beruht der (bedingte) Likelihood-Quotienten-Test für die Modellgeltung? Wie geht man dabei vor?
- Welche beiden Likelihood-Funktionen werden dabei einander gegenübergestellt und wie sind sie definiert?
- Wie lauten Null- und Alternativhypothese des Tests, wie ist die Prüfgröße definiert, und gegen welche Prüfverteilung wird sie getestet?
- Nennen Sie einige Gründe, warum sich die stochastischen Testmodelle in der Psychodiagnostik noch nicht entsprechend durchsetzen konnten!
- Was verstand William Stern unter Intelligenz?
- Stellen Sie Spearmans Generalfaktor-Theorie der Intelligenz und Thurstones Primärfähigkeiten einander gegenüber und erläutern Sie, warum die Kontroverse zwischen Spearman und Thurstone nicht empirisch entschieden werden konnte!
- Welche Primärfähigkeiten wurden von Thurstone postuliert, und durch welche Art von Testleistungen wurden sie operationalisiert?
- Was ist die Faktorenanalyse? Wozu dient sie?
- Was versteht man unter der Stichprobenabhängigkeit faktorenanalytischer Ergebnisse?

- Handelt es sich bei der Intelligenz tatsächlich um eine latente Dimension, auf der man die Probanden anordnen kann?
Falls ja: Welche Gründe lassen sich dafür anführen?
Falls nein: Worum handelt es sich bei der Intelligenz stattdessen, und welche Gründe lassen sich dafür anführen?
- Worum geht es in dem Experiment von Siegler (1976)? Was soll damit gezeigt werden? Welcher Art von Problemlösungsaufgaben bedient es sich, und welche Aufgabentypen werden von Siegler unterschieden?
- Wie beschreibt Siegler die schrittweise Akkumulation von Wissen? Welches sind die vier Entwicklungsstufen, und welche Lösungsstrategien entsprechen ihnen?
- Wie kann die Entwicklungsstufe eines Kindes diagnostiziert werden?
- Warum wäre dies auf Grundlage des Summenscores nicht möglich?
- Wie lauten die prognostizierten Antwortmuster, die entsprechenden stufenspezifischen Lösungswahrscheinlichkeiten und die auf den verschiedenen Entwicklungsstufen erwartete Anzahl richtiger Antworten?
- Geben Sie Beispiele aus Einstellungsdiagnostik, klinischer Psychologie und kognitiver Psychologie, welche dafür sprechen, dass ggf. die Antwortmuster informativer sein können als die Leistungsmenge!
- Erläutern Sie anhand der o.g. Beispiele, was man unter einer qualitativen latenten Variable zu verstehen hat!
- Formulieren Sie ein allgemeines Modell der Latent-Class-Analyse und erläutern Sie die beiden Spezialfälle Ein-Klassen-Lösung und saturiertes Modell!
- Wie kann man die Modellparameter bei $1 < h < h_{\max}$ manifesten Klassen schätzen?
- Wie kann man die Modellparameter
 - a) der Ein-Klassen-Lösung und
 - b) des saturierten Modells
 schätzen? Warum ist dies bei $1 < h < h_{\max}$ latenten Klassen nicht so einfach möglich?
- Stellen Sie den EM-Algorithmus dar und erläutern Sie, warum es sich dabei nicht um ein zirkuläres Vorgehen handelt!
- Was versteht man unter den Membership-Wahrscheinlichkeiten, was beschreiben sie, und wie können sie berechnet werden?
- Erläutern Sie die Anwendung des AIC-Index zur Entscheidung über die Anzahl der latenten Klassen!
- Erläutern Sie, warum man sich zur Auswahl des besten LCA-Modells nicht einfach auf die Betrachtung der Likelihood stützen kann!
- Wie berechnet sich die Anzahl der unabgängigen Modellparameter der LCA
 - a) allgemein,
 - b) im Spezialfall der Ein-Klassen-Lösung,
 - c) im Spezialfall des saturierten Modells?
- Worüber geben Vergleiche des ausgewählten LCA-Modells
 - a) mit der Ein-Klassen-Lösung und
 - b) mit dem saturierten Modell
 Auskunft, und mit welcher Art von statistischem Test können sie durchgeführt werden?
- Wie lautet die Testgröße, und wie ist sie verteilt?

- Welches LCA-Modell entspricht dabei der Nullhypothese, und welches der Alternativhypothese?
- Wie können im Falle der Latent-Class-Analyse diagnostische Entscheidungen über die Vpn gefällt werden?

Übungsaufgaben

1. Das Verhalten der Vpn in einem sozialpsychologischen Experiment wird mittels der Prädikatoren A, B, C und D beschrieben. Die Ergebnisse des Experimentes sind in der untenstehenden Tabelle dargestellt:

Vp-Nr	1	2	3	4	5	6	7	8	9	10	11	12
Verhalten	A	D	C	B	C	D	A	D	B	C	C	C
Zahlenwert	0	3	2	1	1	3	0	4	1	2	2	2

- a) Überprüfen Sie, ob die darunter stehenden Zahlenwerte eine Nominalskala der Verhaltensbeschreibungen bilden, und wenn nein:
 - b) Begründen Sie, warum!
 - c) Bilden Sie eine korrekte Nominalskala der Verhaltensbeschreibungen!
2. Zwischen den Vpn A, B, C und D wurden die folgenden empirischen Relationen beobachtet:

$A \neq B$ $A \neq D$ $B \neq C$ $C \neq C$ $C \neq D$ $D \neq A$ $D \neq B$

- a) Überprüfen Sie, ob die Relation \neq eine Ordnungsrelation ist!
 - b) Wenn nein: Welche Befunde sprechen dagegen?
3. Ein Test aus vier binären Items wurde von 10 Vpn bearbeitet. Die Itemantworten der Vpn sind in der nachstehenden Antwortmatrix wiedergegeben. Überprüfen Sie, ob der Test eine Guttman-Skala bildet, und begründen Sie Ihre Antwort!

Vp-Nr	Item-Nr			
	1	2	3	4
1	0	0	0	0
2	0	1	1	0
3	1	1	1	0
4	0	1	1	1
5	0	1	0	0
6	0	1	0	0
7	1	1	1	0
8	0	0	1	1
9	1	1	1	1
10	1	1	1	0

4. Ein Test aus k dichotomen Items wurde einer Stichprobe von $n = 594$ Vpn zur Lösung vorgegeben. Zwecks Durchführung von Modellgeltungstests wurde die Stichprobe in zwei Teilstichproben zerlegt, in eine Low-Score-Gruppe mit $x_{vo} \leq 2$ und eine High-Score-Gruppe mit $x_{vo} \geq 3$. Nachstehende Tabelle gibt die CML-Itemparameterschätzungen und die bedingte Log-Likelihood der beiden Teilstichproben und der Gesamtstichprobe wieder.

	δ_1	δ_2	δ_3	δ_4	$\ln(\text{CL})$
Low-Score	0.3955	-0.8634	1.4030	-0.9351	-279.8295
High-Score	0.8381	-1.0078	1.5826	-1.4129	-115.9737
Gesamt	0.4966	-0.8987	1.3923	-0.9902	-397.8519

Untersuchen Sie

- mittels graphischer Modellkontrolle und
 - mittels bedingtem Likelihood-Quotienten-Tests, ob die Testleistung der Probanden durch ihre Summenscores erschöpfend beschrieben werden kann!
5. Konstruieren Sie Balkenwaagen-Aufgaben vom Typ
- Balance
 - Gewicht
 - Distanz
 - Konflikt-Gewicht
 - Konflikt-Distanz
 - Konflikt-Balance!
6. Untersuchen Sie, welchem Aufgabentyp (nach Siegler) die folgenden Balkenwaagen-Aufgaben angehören:

Nr.	Links		Rechts	
	G	D	G	D
1	4	2	4	1
2	2	3	1	6
3	4	1	2	3
4	4	6	2	6
5	2	3	4	2
6	5	1	5	1

7. Welche Antworten sind bei den Balkenwaagen-Aufgaben in Aufgabe (6) zu erwarten, wenn sich der Proband (nach Siegler)
- auf Stufe 1
 - auf Stufe 2
 - auf Stufe 3
 - auf Stufe 4
- befindet und die Antworten nicht durch Zufallsschwankungen überlagert werden?
8. Welche Lösungswahrscheinlichkeiten haben die Probanden bei den Balkenwaagen-Aufgaben in Aufgabe (6), wenn sie sich (nach Siegler)

- a) auf Stufe 1
 b) auf Stufe 2
 c) auf Stufe 3
 d) auf Stufe 4
 befinden und die Antworten nicht durch Zufallsschwankungen überlagert werden?
9. Wie viele richtige Antworten auf die Items in Aufgabe (6) sind zu erwarten, wenn sich ein Proband (nach Siegler)
- a) auf Stufe 1
 b) auf Stufe 2
 c) auf Stufe 3
 d) auf Stufe 4
 befindet und die Antworten nicht durch Zufallsschwankungen überlagert werden?
10. Wie unterscheiden sich die Testergebnisse von Probanden der verschiedenen Entwicklungsstufen bei Bearbeitung der Balkenwaagen-Aufgaben in Aufgabe (6)
- a) quantitativ (d. h. hinsichtlich der Lösungsmenge) und
 b) qualitativ (d. h. hinsichtlich der Lösungsmuster)?
11. Wie unterscheiden sich die Testergebnisse von Probanden der Stufen 3 und 4 bei Bearbeitung der Balkenwaagen-Aufgaben in Aufgabe (6)
- a) quantitativ
 b) qualitativ?
12. Eine Vp der Stufe 2 und drei Vpn der Stufe 3 haben die Konflikt-Aufgaben aus Aufgabe (6) bearbeitet. Die Daten seien durch Zufallsfehler überlagert, so dass sich an Stelle der oben berechneten (idealen) Lösungswahrscheinlichkeiten die folgenden Wahrscheinlichkeiten einer richtigen Antwort ergeben:

	Itemtyp		
	K-G	K-D	K-B
	Item-Nummer (neu)		
Stufe	1	2	3
2	0.9	0.1	0.1
3	0.3	0.3	0.3

Die Antworten der Probanden sind in der nachstehenden Antwortmatrix wiedergegeben:

Vp-Nr	Item-Nummer		
	1	2	3
1	1	0	0
2	1	1	0
3	0	1	0
4	0	0	1

- a) Wie groß sind in diesen Daten die Klassengrößen der Stufe 2 (= Klasse 1) und Stufe 3 (= Klasse 2)?

- b) Wie groß ist die Wahrscheinlichkeit, mit welcher
- Vpn der Klasse 1 (= Stufe 2) und
 - Vpn der Klasse 2 (= Stufe 3)
- die beobachteten Antwortmuster (1,0,0), (1,1,0), (0,1,0) und (0,0,1) produzieren?
- c) Wie groß ist die Wahrscheinlichkeit, die oben dargestellte Antwortmatrix zu beobachten?
13. Wie viele verschiedene Antwortmuster sind möglich, wenn ein Test
- aus 5 dichotomen Items,
 - aus 5 dreistufigen Items,
 - aus 3 dichotomen und 2 dreistufigen Items besteht?
14. Mit welcher (Membership-) Wahrscheinlichkeit können die Vpn in Aufgabe (12) den beiden Klassen (= Entwicklungsstufen) zugeordnet werden, und welche diagnostischen Entscheidungen würden Sie treffen?
15. Die Latent-Class-Analyse eines Tests aus 4 dichotomen Items ergab die folgende Goodness-of-Fit-Statistiken:

h	$\ln(\text{ML})$	n(P)	AIC
1	-1523.540	4	3055.080
2	-1347.612	9	2713.223
3	-1332.611	14	2693.222
Sat.	-1331.582	15	

- Entscheiden Sie auf Grundlage der AIC-Indizes, welche Anzahl von latenten Klassen die beste Beschreibung der Daten liefert, und überprüfen Sie mittels des Likelihood-Quotienten-Tests,
- ob das ausgewählte LCA-Modell (Klassenanzahl) die Daten ebenso gut beschreibt wie das saturierte Modell, sowie,
- ob die Verteilung der Daten überhaupt eine Mischverteilung mehrerer latenter Klassen darstellt.

2.5

Experimentelle Hypothesenprüfung

Zufriedenstellen kann eine deduktiv-nomologische Erklärung nur dann, wenn die darin enthaltenen (empirischen) Gesetzesaussagen auch *wahr* sind. D.h. wir müssen sie *überprüfen*. Dazu können wir offensichtlich nur so vorgehen, dass wir die Randbedingungen der Gesetzesaussage(n) variieren und überprüfen, ob die prognostizierte Wirkung auch tatsächlich eintritt. Diese Vorgehensweise bezeichnet man als experimentelle Hypothesenprüfung.

Experimentelle
Hypothesenprüfung

Wir wollen z.B. überprüfen, ob es beim Längenvergleich zweier Linien tatsächlich eine absolute Unterschiedsschwelle gibt. Dazu müssen wir unserem Probanden Paare von verschiedenen langen Linien zeigen und überprüfen, ob es einen kritischen Längenunterschied gibt, unterhalb dessen die Linien nicht mehr als verschieden lang wahrgenommen werden.

Beginnen wir mit einem Standardreiz von $S=20$ cm Länge und variieren dann den Vergleichreiz (V) so, dass die Stimulusdifferenz ($S-V$) 1 mm, 2 mm usw. bis 10 mm beträgt. Die derart vom Experimentator variierte Randbedingung der Gesetzesaussage bezeichnet man als die unabhängige Variable (UV) des Experimentes und unterscheidet sie von der abhängigen Variable (AV), deren Veränderung durch die Gesetzesaussage prognostiziert wird. In unserem Beispiel ist dies die Wahrnehmung der Stimulusdifferenz. Die abhängige Variable entspricht also dem Explanandum des HO-Schemas, und die Gesetzesaussage stellt die AV als Funktion der UV dar.

Unabhängige und
abhängige Variable

$$AV = f(UV) \quad w_{S-V} = \begin{cases} 1 & \text{wenn } |S - V| \geq \Delta S \\ 0 & \text{wenn } |S - V| < \Delta S \end{cases}$$

Die Wahrnehmung des Stimulusunterschieds haben wir auf einer binären Nominalskala gemessen. Die Skalenpunkte sind $w_{S-V} = 1$, wenn der Unterschied wahrgenommen wird ($w(S-V)$), und

$w_{S-V} = 0$, wenn der Unterschied nicht wahrgenommen wird ($-w(S-V)$).

Zum Beispiel erhalten wir die in Tab. 2.5.1 dargestellten Ergebnisse, die mit der postulierten empirischen Gesetzmäßigkeit übereinstimmen:

Tab. 2.5.1: Mögliche Ergebnisse eines Experimentes zur Identifikation der absoluten Unterschiedsschwelle

S	UV	AV
	S-V	w(S-V)
20 cm	1.0 mm	0
	2.0 mm	0
	3.0 mm	0
	4.0 mm	1
	5.0 mm	1
	6.0 mm	1
	7.0 mm	1
	8.0 mm	1
	9.0 mm	1
	10.0 mm	1

Die Längenunterschiede bis einschließlich 3 mm wurden nicht wahrgenommen, die Längenunterschiede von 4 mm und darüber wurden wahrgenommen. Es gibt also in diesem Experiment eine absolute Unterschiedsschwelle.

Diese liegt im Intervall $3 \text{ mm} < \Delta S \leq 4 \text{ mm}$. Die relative Unterschiedsschwelle $c = \Delta S/S$ liegt somit zwischen $1/66.7$ und $1/50$.

Die Hypothesenförmigkeit empirischer Gesetzaussagen

Unsere Gesetzaussage haben wir damit aber trotzdem noch nicht bewiesen. Denn wir haben die Existenz einer absoluten Unterschiedsschwelle bisher nur für eine bestimmte V_p , für einen bestimmten Standardreiz $S=20 \text{ cm}$ und für eine Auswahl von Vergleichsreizen gezeigt, die zudem alle kürzer waren als der Standardreiz. Unsere Gesetzaussage behauptet die Existenz einer absoluten Unterschiedsschwelle aber für alle möglichen Stimuluspaare – und auch nicht nur für diese eine V_p , sondern für alle möglichen Probanden. Um unsere Gesetzaussage zu überprüfen, müssen wir das Experiment noch sowohl mit anderen Standard- und Vergleichsreizen als auch mit anderen Probanden durchführen.

Selbst wenn wir dabei jedes Mal Ergebnisse erzielen, die – wie die obigen – mit der Gesetzaussage übereinstimmen, werden wir

die Geltung der Gesetzesaussage niemals definitiv bewiesen haben. Selbst wenn wir noch so viele Experimente durchgeführt haben, können wir nie wissen, ob es nicht doch eine Vp oder ein Stimuluspaar gibt, für die das Gesetz nicht gilt: Aufgrund ihrer logischen Struktur als universelle empirische Subjunktionen können empirische Gesetzesaussagen niemals definitiv bewiesen werden, sondern sie bleiben stets Hypothesen (= Vermutungen). Als solche können sie beibehalten werden, bis sich ein Gegenbeispiel findet, welches die Hypothese falsifiziert. Die Verwerfung der Gesetzhypothese aufgrund eines Gegenbeispiels folgt dabei der logischen Schlussform

$$(A \wedge \neg B) \rightarrow \neg(A \rightarrow B).$$

Zusammenfassend können wir sagen, dass Naturgesetze aufgrund ihrer logischen Form nicht verifizierbar, sondern nur falsifizierbar sind. Sie bleiben stets Hypothesen, die mehr oder minder gut bewährt sind, indem sie möglichst kritischen und systematischen Falsifikationsversuchen standgehalten haben (vgl. Popper, 1994). Als Instrument der Geltungsprüfung dient das Experiment, in dem die Randbedingungen der Gesetzesaussage als unabhängige Variable (UV) systematisch variiert und ihre Wirkungen als abhängige Variable (AV) beobachtet bzw. gemessen werden. Stimmt der Ausgang des Experimentes mit der postulierten Gesetzmäßigkeit $AV = f(UV)$ überein, so sagt man, die Hypothese habe sich bestätigt oder bewährt. Andernfalls ist die Hypothese falsifiziert.

Falsifikation und
Bewährung

Dabei kann zwischen verschiedenen Graden der Bewährung unterschieden werden. Eine Hypothese, die oft geprüft und dabei nie falsifiziert wurde, gilt als gut bewährt. Eine neue Hypothese, die erst selten geprüft wurde, ist noch nicht so gut bewährt usw.

2.5.1 Grundbegriffe des Experimentierens

Die verschiedenen Ausprägungen einer Variable, die im Experiment realisiert werden, bezeichnet man als die Variablenstufen. In dem oben dargestellten Experiment zur Feststellung der absoluten Unterschiedsschwelle hatte z.B. die unabhängige Variable (S-V) die Stufen 1mm, 2mm ... 10mm. Die abhängige Variable w_{S-V} hatte die Stufen 1 (= $w(S-V)$) und 0 (= $\neg w(S-V)$).

Variablenstufen

Einfaktorielle und mehrfaktorielle Experimente

In diesem Falle handelte es sich um ein einfaktorielles Experiment, in dem nur eine UV (die Stimulusdifferenz $S-V$) variiert wurde. Experimente, in denen mehrere UVn gleichzeitig im Spiel sind, bezeichnet man als mehrfaktorielle Experimente und nennt die verschiedenen UVn auch die (experimentellen) Faktoren.

Die Stufen der UV (bzw. bei mehrfaktoriellen Versuchsplänen: die Kombinationen der verschiedenen Faktorstufen) bezeichnet man als die Versuchsbedingungen oder die experimentellen Bedingungen.

Ein Versuchsplan zur Überprüfung des Weber'schen Gesetzes

Unseren Versuchsplan zum Nachweis der absoluten Unterschiedsschwelle könnten wir z.B. zu einem dreifaktoriellen Versuchsplan ausweiten, indem wir auch die Größe des Standardreizes S und das Vorzeichen der Stimulusdifferenz $S-V$ variieren. Gegenüber unserem bisherigen (einfaktoriellen) Versuchsplan hätte dies den Vorteil, dass sich im Falle der Bestätigung der Hypothese zugleich auch ihr Bewährungsgrad deutlich erhöhen würde: Wir hätten die Hypothese ja dann nicht nur für einen Standardreiz ($S=20$ cm) und für positive Stimulusdifferenzen von 1 mm bis 10 mm Länge überprüft, sondern für mehrere Standardreize und auch für negative Stimulusdifferenzen. Z.B. wählen wir

- für den Standardreiz die Faktorstufen 10, 20, 30 und 40 cm,
- für den Betrag der Stimulusdifferenz $|S-V|$ (wie gehabt) die Faktorstufen $S/200$, $2 \times S/200$, $3 \times S/200$... $10 \times S/200$, und
- für das Vorzeichen der Stimulusdifferenz die Faktorstufen $S-V < 0$ und $S-V > 0$,

und erhalten die in Tab. 2.5.2 dargestellten $4 \times 10 \times 2 = 80$ Versuchsbedingungen. Bei der abhängigen Variable unterscheiden wir wieder die beiden Variablenstufen 1 (= $w(S-V)$) und 0 (= $-w(S-V)$).

Tab. 2.5.2: Ein $4 \times 10 \times 2$ -Versuchsplan zur Überprüfung des Weber'schen Gesetzes. Vergleichsreize in Millimetern

S	S-V	S-V	
		< 0	> 0
10 cm	$S/200$	99.5 mm	100.5 mm
	$2 \times S/200$	99.0 mm	101.0 mm
	$3 \times S/200$	98.5 mm	101.5 mm
	$4 \times S/200$	98.0 mm	102.0 mm
	$5 \times S/200$	97.5 mm	102.5 mm
	$6 \times S/200$	97.0 mm	103.0 mm

	7×S/200	96.5 mm	103.5 mm
	8×S/200	96.0 mm	104.0 mm
	9×S/200	95.5 mm	104.5 mm
	10×S/200	95.0 mm	105.0 mm
20 cm	S/200	199.0 mm	201.0 mm
	2×S/200	198.0 mm	202.0 mm
	3×S/200	197.0 mm	203.0 mm
	4×S/200	196.0 mm	204.0 mm
	5×S/200	195.0 mm	205.0 mm
	6×S/200	194.0 mm	206.0 mm
	7×S/200	193.0 mm	207.0 mm
	8×S/200	192.0 mm	208.0 mm
	9×S/200	191.0 mm	209.0 mm
	10×S/200	190.0 mm	210.0 mm
30 cm	S/200	298.5 mm	301.5 mm
	2×S/200	297.0 mm	303.0 mm
	3×S/200	295.5 mm	304.5 mm
	4×S/200	294.0 mm	306.0 mm
	5×S/200	292.5 mm	307.5 mm
	6×S/200	291.0 mm	309.0 mm
	7×S/200	289.5 mm	310.5 mm
	8×S/200	288.0 mm	312.0 mm
	9×S/200	256.5 mm	313.5 mm
	10×S/200	285.0 mm	315.0 mm
40 cm	S/200	398.0 mm	402.0 mm
	2×S/200	396.0 mm	404.0 mm
	3×S/200	394.0 mm	406.0 mm
	4×S/200	392.0 mm	408.0 mm
	5×S/200	390.0 mm	410.0 mm
	6×S/200	388.0 mm	412.0 mm
	7×S/200	386.0 mm	414.0 mm
	8×S/200	384.0 mm	416.0 mm
	9×S/200	382.0 mm	418.0 mm
	10×S/200	380.0 mm	420.0 mm

Hypothesen und empirische Prognose

Nachdem wir bereits oben empirisch festgestellt hatten, dass die relative Unterschiedsschwelle zwischen $1/50$ bis $1/66.7$ lag, können wir diesen Versuchsplan benutzen, um nicht nur die Existenz einer absoluten Unterschiedsschwelle, sondern auch das Weber'sche Gesetz zu überprüfen, und sogar noch mehr. Ein und derselbe Versuchsplan erlaubt es uns nämlich jetzt, drei Hypothesen gleichzeitig zu prüfen:

1. Dass es für alle Standardreize eine absolute Unterschiedsschwelle gibt, so dass:

$$\forall_s \cdot w_{s-v} = \begin{cases} 1 & \text{wenn } |S - V| \geq \Delta S \\ 0 & \text{wenn } |S - V| < \Delta S \end{cases}$$

2. Dass die absolute Unterschiedsschwelle proportional zum Standardreiz ist:

$$\forall_s \cdot \Delta S = c \times S$$

3. Dass die relative Unterschiedsschwelle beim Längenvergleich zwischen $1/66.7$ und $1/50$ liegt:

$$\forall_s \cdot (1/66.7 < c \leq 1/50)$$

Treffen diese drei Hypothesen zu, so sieht das erwartete Ergebnis unseres Experimentes aus wie in Tab. 2.5.3 dargestellt.

Tab. 2.5.3: Erwartetes Ergebnis eines $4 \times 10 \times 2$ Experimentes zur Überprüfung des Weber'schen Gesetzes

S	S-V	S-V	
		< 0	> 0
10 cm	S/200	0	0
	2×S/200	0	0
	3×S/200	0	0
	4×S/200	1	1
	5×S/200	1	1
	6×S/200	1	1
	7×S/200	1	1
	8×S/200	1	1
	9×S/200	1	1
	10×S/200	1	1

20 cm	$S/200$	0	0
	$2 \times S/200$	0	0
	$3 \times S/200$	0	0
	$4 \times S/200$	1	1
	$5 \times S/200$	1	1
	$6 \times S/200$	1	1
	$7 \times S/200$	1	1
	$8 \times S/200$	1	1
	$9 \times S/200$	1	1
	$10 \times S/200$	1	1
30 cm	$S/200$	0	0
	$2 \times S/200$	0	0
	$3 \times S/200$	0	0
	$4 \times S/200$	1	1
	$5 \times S/200$	1	1
	$6 \times S/200$	1	1
	$7 \times S/200$	1	1
	$8 \times S/200$	1	1
	$9 \times S/200$	1	1
	$10 \times S/200$	1	1
40 cm	$S/200$	0	0
	$2 \times S/200$	0	0
	$3 \times S/200$	0	0
	$4 \times S/200$	1	1
	$5 \times S/200$	1	1
	$6 \times S/200$	1	1
	$7 \times S/200$	1	1
	$8 \times S/200$	1	1
	$9 \times S/200$	1	1
	$10 \times S/200$	1	1

Entspricht das Ergebnis des Experimentes dieser Prognose, so haben sich unsere Hypothesen bewährt. Stimmt das Ergebnis *nicht* mit der Prognose überein, so können eine oder auch mehrere der Hypothesen falsch sein. Anhaltspunkte dafür, *welche* der Hypothesen falsch sind und wie sie ggf. modifiziert werden müs-

Abweichende
Ergebnisse

sen, können wir dann mittels einer genaueren Inspektion der experimentellen Daten beziehen. Z.B. mögen die abweichenden Ergebnisse nur bei dem Standardreiz von 10 cm Länge aufgetreten sein und wie in Tab. 2.5.4 aussehen.

Tab. 2.5.4: Von der Prognose abweichende Ergebnisse eines Experimentes zur Überprüfung des Weber'schen Gesetzes (grau unterlegt)

S	S-V	S-V	
		< 0	> 0
10 cm	S/200	0	0
	2×S/200	0	0
	3×S/200	0	0
	4×S/200	0	0
	5×S/200	0	0
	6×S/200	1	1
	7×S/200	1	1
	8×S/200	1	1
	9×S/200	1	1
	10×S/200	1	1

Dann könnten wir auf Grundlage dieser Ergebnisse vermuten,

- dass die relative Unterschiedsschwelle von $1/66.7$ bis $1/50$ nur für Standardreize zwischen 20 und 40 cm Länge gilt und
- dass die relative Unterschiedsschwelle für kleinere Standardreize (um 10 cm Länge) jedoch gröber ist und zwischen $1/40$ und $1/33.3$ liegt.

Der Wechsel von Induktion und Deduktion

Diese Verallgemeinerung (Generalisierung) der im Experiment gewonnenen Ergebnisse beruht jedoch auf einem induktiven Schluss, der – im Unterschied zu den deduktiven Schlussformen des modus ponens und des modus tollens – logisch nicht begründbar ist.

Deduktiv – und damit logisch zwingend – ist

- aufgrund des modus ponens: die Erklärung des Einzelfalls aus der Hypothese und den Randbedingungen, und damit auch
- die Prognose des Ausgangs des Experimentes, sowie
- aufgrund der Schlussform $(A \wedge \neg B) \rightarrow \neg(A \rightarrow B)$: die Falsifikation der Hypothese aufgrund abweichender Ergebnisse.

Bei der Generalisierung unserer Beobachtungen schließen wir dagegen induktiv vom Einzelfall auf eine allgemeine Gesetzmäßig-

keit. Da dieser Schluss nicht logisch zwingend ist, müssen wir die induktiv gewonnene neue Hypothese erneut empirisch (bzw. experimentell) überprüfen. Der Forschungsprozess in den Erfahrungswissenschaften stellt sich folglich als ein ständiger Wechsel zwischen induktiven und deduktiven Vorgehensweisen dar (vgl. Abb. 2.5.1).

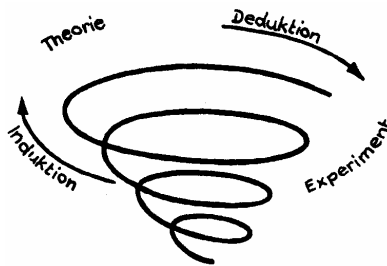


Abb. 2.5.1: Ein Modell für die Beziehungen des Experimentes zu den Denkmethoden (nach Selg, 1975, 12)

Zur Prüfung der neuen Hypothese können wir aber nicht wieder auf die alten Daten zurückgreifen, denn wir haben die Hypothese ja aufgrund dieser Daten erstellt. D.h. wir haben sie von vornherein so formuliert, dass sie mit unseren bisherigen Daten übereinstimmt. Um die Hypothese überprüfen zu können – d.h. um ihr überhaupt eine Chance zu geben, an der Erfahrung zu scheitern –, müssen wir (mit veränderter Prognose) das Experiment wiederholen oder ein neues Experiment durchführen.

Die Notwendigkeit der Hypothesenformulierung ex ante

Zusammenfassend können wir festhalten:

- Stimmt der Ausgang des Experimentes nicht mit der Prognose überein, so ist mindestens eine der dem Experiment zugrundeliegenden Hypothesen falsch.
- Deshalb wird man in der Regel aber nicht gleich die ganze Theorie verwerfen, sondern die Theorie auf der Grundlage der Daten modifizieren.
- Die neue (bzw. modifizierte) Theorie ist so zu formulieren, dass sie mit den bisherigen Daten übereinstimmt (andernfalls ist sie von vornherein falsch).
- Zu ihrer Überprüfung bedarf es eines neuerlichen Experimentes.

Allgemein lässt sich die Regel aufstellen, dass die zu prüfenden Hypothesen bereits vor der Erhebung der Daten bzw. vor der Durchführung des Experimentes formuliert sein müssen. Diese methodische Reihenfolge (Hypothese vor der Prüfung) ist nicht

nur dafür erforderlich, dass man das Experiment geeignet planen kann, sondern auch, dass man überhaupt sinnvoller Weise von der Prüfung der Hypothese sprechen kann.

Störvariablen

Auch im Falle von abweichenden Ergebnissen werden wir unsere Theorie nicht in jedem Fall gleich verwerfen bzw. modifizieren. Es könnte nämlich auch sein, dass wir schlichtweg bei der Planung und Durchführung unseres Experimentes einen Fehler begangen haben, indem wir z.B. andere Variablen übersehen haben, welche den Ausgang des Experimentes ebenfalls beeinflussen können, die in unserer Theorie aber nicht vorkommen bzw. bei der Durchführung des Experimentes nicht kontrolliert wurden. Solche Variablen, welche die AV beeinflussen können und deren Einfluss ausgeschaltet werden muss, wenn das Experiment schlüssige Aussagen über die Geltung der Theorie zulassen soll, bezeichnet man als Störvariablen (SV).

Für die Kontrolle von Störvariablen gibt es verschiedene Techniken, auf die wir im Kapitel über "Das psychologische Experiment" in Band II unserer Einführung zu sprechen kommen werden. Trotz aller Kontrolltechniken kann es aber passieren, dass eine SV mit der UV systematisch variiert. Man sagt dann, dass die SV mit der UV konfundiert ist. Das Ergebnis des Experimentes kann dann nicht mehr auf die UV zurückgeführt werden, da die Wirkung der UV nicht von der Wirkung der SV getrennt werden kann.

Z.B. ist für die relative Unterschiedsschwelle, die beim Längenvergleich zwischen $1/50$ und $1/66.7$ des Standardreizes liegen soll, die Normalsichtigkeit der Probanden vorausgesetzt bzw. dass Fehlsichtigkeit während des Experimentes durch eine Sehhilfe (Brille) korrigiert wird. Da wir die Sehschärfe der Probanden in dem oben beschriebenen Experiment nicht kontrolliert haben, könnte es sein, dass die V_p , an der wir das Experiment mit dem Standardreiz von 10 cm durchgeführt haben, kurz- oder weitsichtig war, und dass die Unterschiedsschwelle bei dieser V_p aufgrund ihrer Fehlsichtigkeit heraufgesetzt war.

Exhaustion

Sollte dies der Fall sein, so werden wir unsere ursprüngliche Hypothese, wonach die Unterschiedsschwelle auch bei einem Standardreiz von 10 cm Länge zwischen $1/50$ und $1/66.7$ des Standardreizes liegt, trotz abweichender Befunde beibehalten. Eine solche Verteidigung einer Hypothese gegen die empirischen Daten bezeichnet man als Exhaustion. Welche Art von Argumenten für die Exhaustion einer Hypothese vorgebracht werden können, wird

in Band II ausführlicher diskutiert werden. Grob einteilen kann man sie nach Campbell & Stanley (1963) in mangelnde interne und/oder mangelnde externe Validität (= Gültigkeit) des Experimentes.

- Die interne Validität eines Experimentes ist gegeben, wenn im Experiment nur zum Tragen kommt, was auch in der Theorie (bzw. in der Hypothese) vorkommt. Das bedeutet im Idealfall die Kontrolle aller möglichen Störvariablen und erfordert in der Regel Laborexperimente, die in einer mehr oder minder künstlichen Untersuchungssituation, meist in einem speziellen Untersuchungsraum an einem Forschungsinstitut, durchgeführt werden.
- Die Optimierung der internen Validität eines Experimentes geht jedoch auf Kosten seiner Lebensnähe, so dass nicht ohne weiteres sichergestellt werden kann, dass die empirischen Forschungsergebnisse auch unter natürlichen Bedingungen gelten, d.h. ob sie auch externe Validität besitzen.

Interne und externe Validität

Bei so einfachen psychophysischen Experimenten, wie wir sie bisher (z.B. zum Weber'schen Gesetz) kennen gelernt haben, spielen Zweifel an der externen Validität des Experimentes kaum eine Rolle. Ganz anders sieht es dagegen z.B. in der Sozialpsychologie aus.

Den höchsten Grad an externer Validität kann man mit Hilfe von Feldstudien erreichen, die in der natürlichen Lebensumwelt der Probanden durchgeführt werden, ohne dass überhaupt ein experimenteller Eingriff in die Untersuchungssituation stattfindet. Indem dabei eine Vielzahl an nicht kontrollierbaren SVn ins Spiel kommen können, geht dies aber auf Kosten der internen Validität der Untersuchung.

Einen Kompromiss zwischen interner und externer Validität schließen die Feldexperimente, die ebenfalls in der natürlichen Lebensumwelt der Vpn durchgeführt werden, bei denen sich der Forscher aber nicht auf die Rolle des unbeteiligten Beobachters beschränkt, sondern – wie im Laborexperiment auch – gezielte Eingriffe in die Untersuchungssituation vornimmt (vgl. Tab. 2.5.5).

	Vorteil	Nachteil
Laborexperimente	Hohe interne Validität	Geringe externe Validität
Feldexperimente	Kompromiss zwischen interner und externer Validität	
Feldstudien	Hohe externe Validität	Geringe interne Validität

Tab. 2.5.5: Laborexperimente, Feldstudien und Feldexperimente

2.5.1.1 Exkurs: Das Ferienlagerexperiment von Sherif

Ein Beispiel hierfür ist das berühmte Ferienlager-Experiment von Muzafer Sherif (Sherif & Sherif, 1969), mit dem gezeigt werden sollte, wie sich Kooperation und Konkurrenz auf das Verhalten zwischen Gruppen auswirken. Dazu wählte Sherif eine Versuchsanordnung, die nicht im Labor, sondern in realen Lebenssituationen stattfand. Sherif führte seine Experimente auf einem eigens dafür eingerichteten Lagerplatz durch. Seinen Versuchspersonen, Jungen im Alter von elf bis zwölf Jahren, war nicht bekannt, dass ihr Verhalten unter Beobachtung stand. Das Experiment gliederte sich in drei Abschnitte.

Gruppenbildung	Der erste Abschnitt dauerte etwa eine Woche und diente dem Prozess der Gruppenbildung. Er begann damit, dass jede der beiden Gruppen in einem eigenen Autobus ins Lager kam und in einer Unterkunft wohnte. Jede der beiden Gruppen war mit einer Vielzahl von kooperativen Tätigkeiten beschäftigt: Kampieren im Freien, Kochen, Verbesserung von Schwimmgelegenheiten, Transport von Booten über raues Gelände ins Wasser usw. Während dieser Zeit entwickelten die Gruppen ihre eigenen Organisationsformen, Anzeichen eines Wir-Gefühls und eines Gefühls für gemeinsame Errungenschaften.
Konflikteskalation	Im zweiten Abschnitt des Experimentes wurden durch Konkurrenzreize Spannungen zwischen den beiden Gruppen erzeugt. Sherif schuf eine Reihe von Situationen (z.B. Wettbewerbsspiele), in denen die eine Gruppe ihr Ziel nur auf Kosten der anderen erreichen konnte. Während dieser Phase entwickelten die Gruppenmitglieder feindselige Einstellungen und Klischeehaltungen gegenüber der anderen Gruppe und ihren Mitgliedern. Der Konflikt offenbarte sich schließlich in abwertenden Schimpfwörtern und Schmährufen, im Aufflammen körperlicher Gewalt und in Angriffen auf die Unterkunft der anderen. Gleichzeitig nahmen das Zusammengehörigkeitsgefühl und die Hilfsbereitschaft innerhalb der jeweils eigenen Gruppe zu, und es vollzog sich ein Wandel in der sozialen Rangordnung der Gruppenmitglieder, wobei diejenigen, die sich in der Konkurrenz zwischen den Gruppen besonders hervortaten, im Rang stiegen.
Konfliktreduktion	Im dritten Abschnitt des Experimentes wurden Maßnahmen ergriffen, die den Konflikt zwischen den Gruppen reduzieren sollten. Dazu gehörten Kontakte zwischen den Gruppen bei Handlungen,

die für die Gruppenmitglieder zwar angenehm waren, aber keinerlei Abhängigkeit zwischen den Gruppen implizierten, wie z.B. gemeinsam ins Kino zu gehen, im gleichen Speisesaal zu essen usw. Eine andere Maßnahme war die Einführung übergeordneter Ziele (z.B. die Reparatur der Wasserversorgung auf dem Zeltplatz, die Beschaffung von Geld für den Besuch eines sehenswerten Films oder die Instandsetzung eines zusammengebrochenen Lastwagens), welche von keiner Gruppe ohne Mitarbeit der anderen erreicht werden konnten.

Die Ergebnisse zeigten, dass intergruppale Kontakte ohne übergeordnete Ziele nicht zur Heilung der Beziehung zwischen den Gruppen beitrugen, sondern im Gegenteil oft in Streit ausarteten. Die gemeinsame Anstrengung, gruppenübergreifende Ziele zu erreichen, veränderte dagegen allmählich die Beziehung zwischen den beiden Gruppen. Feindschaft wandelte sich in wohlwollende Interaktion.

Was Sherifs Experiment so spannend macht, ist unter anderem, dass es sich dabei um ein lebensnahes Feldexperiment handelt, das in einer komplexen sozialen Situation durchgeführt wurde. Das Ferienlagerexperiment unterscheidet sich aber auch in anderer Hinsicht von den zuvor dargestellten Laborexperimenten zum Weber'schen Gesetz:

Formale
Eigenschaften des
Experimentes

1. Es handelt sich um ein multivariates Experiment, in dem eine Vielzahl an AVn beobachtet wurden (Wir-Gefühl innerhalb der eigenen Gruppe [in-group], feindselige Einstellungen, verbale und körperliche Gewalt gegenüber der Außengruppe [out-group] usw.). Experimente mit nur einer AV bezeichnet man dagegen als univariat.
2. Dem Experiment liegt ein Versuchsplan zugrunde, in dem dieselben Vpn unter drei verschiedenen experimentellen Bedingungen (Gruppenbildung, Konkurrenz und Kooperation) beobachtet wurden. Solche Versuchspläne bezeichnet man als Versuchspläne mit abhängigen Stichproben und unterscheidet sie von Versuchsplänen mit unabhängigen Stichproben, bei denen jeder Proband nur unter einer einzigen Versuchsbedingung beobachtet wird, so dass jeder Versuchsbedingung eine eigene Versuchsgruppe entspricht.

Univariate und
multivariate
Versuchspläne

Abhängige und
unabhängige
Stichproben

Bei einfaktoriellen Versuchsplänen, in denen nur zwei Variablenstufen der UV unterschieden sind, werden die beiden experimentellen Bedingungen oft auch als Versuchsbedingung (VB) und

Versuchsgruppe und
Kontrollgruppe

Kontrollbedingung (KB) bezeichnet. Entsprechend nennt man die beiden Gruppen von Probanden, welche diesen Bedingungen ausgesetzt werden, die Versuchsgruppe und die Kontrollgruppe.

2.5.1.2 Exkurs: Das Milgram-Experiment

Was dem Ferienlagerexperiment so hohe externe Validität verleiht, ist nicht nur die Tatsache, dass es in einer lebensnahen Versuchssituation durchgeführt wurde. Sherif ist es vielmehr auch gelungen, wesentliche Strukturmerkmale des Gruppenprozesses im Experiment nachzubilden. Eine solche Nachbildung relevanter Situationsmerkmale im Experiment kann ggf. auch in Laborexperimenten gelingen, wie die Experimente zur Gehorsamsbereitschaft von Milgram (1974) zeigen.

Die Versuchs-
anordnung

Um die Gehorsamsbereitschaft zu studieren, entwarf Milgram eine experimentelle Anordnung, in der zwei Leute ein Psychologie-Labor betreten, um an einer Untersuchung über Erinnerungsvermögen und Lernfähigkeit teilzunehmen. Einer von ihnen wird zum "Lehrer" bestimmt, der andere zum "Schüler". Der Versuchsleiter erklärt ihnen, dass sich die Untersuchung mit den Auswirkungen von Strafe auf das Lernen befasst.

Tatsächlich ist nur der Lehrer Versuchsperson. Er hat die Aufgabe, dem Schüler, einem instruierten Versuchsleitergehilfen, der im Nebenraum sitzt, eine festgelegte Reihe von Aufgaben zu stellen. Bei richtiger Beantwortung wird in der Liste fortgefahren. Bei falschen Antworten soll ein Elektroschock in steigender Höhe, beginnend mit 15 Volt, gegeben werden. Ab einer gewissen Stromstärke beginnt der Schüler (der seine Rolle nur spielt und in Wirklichkeit natürlich keine Schocks erhält) Unbehagen zu äußern. Bei 75 Volt murrte er, bei 120 Volt beklagt er sich ausdrücklich, bei 150 Volt bittet er darum, aus dem Experiment entlassen zu werden. Die Proteste steigern sich, werden heftiger und stärker emotional gefärbt. Bei 285 Volt kann die Reaktion nur noch als qualvolles Schreien bezeichnet werden. Möchte der Lehrer das Experiment beenden, so wird er vom Versuchsleiter nachdrücklich aufgefordert, damit fortzufahren.

"Für die Versuchsperson ist die gegebene Situation kein Spiel; ihr Konflikt ist heftig und deutlich erkennbar. Einerseits drängt die offenkundige Qual des Schülers sie dazu, die Sache aufzugeben. Andererseits befiehlt ihr der Versuchsleiter – also eine legitime Autorität, der sie sich in gewisser Weise verpflichtet fühlt –, das Experiment fortzusetzen" (Milgram, 1974, 20).

In den ersten 40 Experimenten, an denen jeweils 40 männliche Versuchspersonen verschiedener Alters- und Berufsgruppen teilnahmen, wurde die Nähe zum Opfer variiert. Die Ergebnisse dieses Experimentes und seiner Variationen erschütterten die Öffentlichkeit:

Einige Ergebnisse

- Nur 35 Prozent der Versuchspersonen widersetzten sich dem Versuchsleiter bei Fernraum-Anordnung ohne akustische Rückmeldung durch das Opfer,
- 37.5 Prozent bei der soeben beschriebenen akustischen Rückmeldung durch das Opfer,
- 60 Prozent im Raumnähe-Versuch und
- 70 Prozent bei Berührungsnähe.

Umgekehrt ausgedrückt: Fast ein Drittel der Versuchspersonen war in der 4. Versuchsbedingung selbst dann noch bereit, mit dem Experiment fortzufahren, wenn der Versuchsleiter ihnen befahl, die Hand des Schülers mit Gewalt auf die Schockplatte zu drücken.

Wiederholungen des Experimentes in anderen Ländern zeigten prinzipiell ähnliche Ergebnisse. Diese Zahlen sind dramatisch und haben Aufsehen erregt. Das wirklich interessante an dem Milgram-Experiment sind aber nicht diese Zahlen, sondern die Selbstauskünfte der Vpn, die anschließend interviewt wurden.

Selbstauskünfte der Vpn

Darin berichteten die Vpn z.B. über ihren inneren Zwiespalt:

"Wenn ich jetzt aufhöre, war all das Leid, das ich dem ‚Schüler‘ bereits zugefügt habe, sinnlos, die Daten aus dem Experiment sind dann nicht verwertbar, das Experiment muss mit einem anderen Lehrer-Schüler-Paar wiederholt werden. Dann muss noch einmal jemandem Leid antan werden". "Wenn ich weitermache, muss ich den ‚Schüler‘ noch weiter quälen, aber die Qual ist wenigstens nicht sinnlos und dient dem Fortschritt der Wissenschaft" usw.

Durch diese Versuchsanordnung ist es Milgram gelungen, in seinen Probanden einen Konflikt zwischen zwei Intentionen zu induzieren, für dessen Realitätshaltigkeit die Laborsituation selbst ein wesentliches Strukturmerkmal darstellte: einen Konflikt zwischen dem Gehorsam gegenüber dem VL und der Institution ‚Wissenschaft‘ einerseits und andererseits dem Wunsch, den Schüler nicht weiter quälen zu müssen.

Der innere Konflikt der Probanden

Auf der Grundlage seiner Ergebnisse (vor allem der Interviews mit den Probanden) hat Milgram (1974) eine umfassende Theorie

des Gehorsams entwickelt und in Form eines kybernetischen Modells (= Regelkreismodell) dargestellt, welches die Veränderungen beschreibt, die eintreten, wenn das autonom handelnde Individuum in einer hierarchischen Sozialstruktur verankert wird, in der es nicht mehr selbständig, sondern als Komponente des Systems fungiert.

Prüfexperimente
und Erkundungs-
experimente

Damit wird deutlich, dass es sich beim Milgram-Experiment um etwas ganz anderes handelt als etwa bei unserem Experiment zur Überprüfung des Weber'schen Gesetzes.

Das Milgram-Experiment ist nicht erster Linie ein Prüfelement, das dazu dient, den Wahrheitsgehalt vorgefasster Hypothesen zu überprüfen (obwohl es – z.B. mit der Untersuchung von Geschlechtsunterschieden bezüglich der Gehorsamsbereitschaft – auch solche Aspekte enthält). Es ist aber auch kein bloßes Erkundungsexperiment, das dazu dient, die empirische Wirklichkeit zu erkunden, um eine Hypothese im Sinne des deduktiv-nomologischen Erklärungsmodells aufstellen zu können.¹

Jenseits des
deduktiv-
nomologischen
Erklärungsmodells

Das Milgram-Experiment stellt vielmehr die prototypische Nachbildung einer bestimmten sozialen Situation dar, mit dem Ziel, die subjektive Wirklichkeit der Menschen, die in dieser Situation verfangen sind, exemplarisch studieren zu können. Damit bewegt sich Milgram auch nicht mehr im deduktiv-nomologischen Erklärungsmodell, sondern im narrativen Erklärungsmodell, auf das wir weiter unten eingehen werden.

Funktionen des
Experimentes

Hier sei nur so viel festgehalten, dass zwar einerseits das deduktiv-nomologische Erklärungsmodell das Experiment als Methode zur Geltungsprüfung von Theorien nach sich zieht; dass dies aber andererseits nicht die einzige Funktion ist, welche experimentelle Forschung erfüllen kann. Entsprechend sind auch die Kriterien, nach denen ein Experiment aufzubauen und zu bewerten ist, nicht für alle Experimente in gleicher Weise schematisch fest-schreibbar. Das Experiment dient im Kontext der intentionalen und/oder der narrativen Erklärung einer anderen Funktion als im Rahmen des deduktiv-nomologischen und/oder des induktiv-sta-tistischen Erklärungsmodells. Und je nach Erklärungsmodell und

1. Empirische Untersuchungen mit Erkundungsabsicht bezeichnet man allgemein als Pilot-Studien und unterscheidet sie von Voruntersuchungen, die dazu dienen, einen Untersuchungsplan, die verwendeten Messinstrumente etc. vor Durchführung der eigentlichen Untersuchung an einer kleinen Stichprobe von Probanden zu erproben (und ggf. zu verbessern).

Funktion des Experimentes im Rahmen dieses Modells werden verschiedene Qualitätsmaßstäbe anzulegen sein.

2.5.2 Voraussetzungen experimenteller Hypothesenprüfung

Naturwissenschaftliche Theorien, die nach dem Schema der deduktiv-nomologischen Erklärung aufgebaut sind, bestehen in der Regel nicht bloß aus einer Gesetzesaussage, sondern sie sind mehr oder minder komplexe Systeme von Gesetzesaussagen, die miteinander strukturell verbunden sind. Indem die verschiedenen Naturgesetze gleichzeitig wirken, ist es dabei auch nicht möglich, eine einzelne Gesetzesaussage isoliert von den anderen und für sich allein zu überprüfen. Zur Überprüfung solcher Theorien kann man nur so vorgehen, dass man aus der Theorie eine konkrete Sachhypothese ableitet und diese dann experimentell untersucht.

Sachhypothesen

Damit der Ausgang des Experimentes möglichst aussagekräftig bezüglich der Geltung der Theorie ist, wird man bevorzugt solche Sachhypothesen untersuchen, die einen möglichst hohen empirischen Gehalt haben, und man wird das Experiment so gestalten, dass es eine möglichst strenge Theorienprüfung erlaubt.

Der empirische Gehalt einer Hypothese ist um so größer, je größer die Menge ihrer Falsifikationsmöglichkeiten ist.

Empirischer Gehalt

So könnten wir z.B. zwecks Überprüfung des Weber'schen Gesetzes die folgenden Sachhypothesen formulieren, die sich in ihrem empirischen Gehalt deutlich voneinander unterscheiden:

1. Längenunterschiede zwischen einem Standardreiz von $S=10$ cm und einem Vergleichsreiz $V \neq S$ werden wahrgenommen oder nicht.
2. Bei einem Standardreiz von $S=10$ cm gibt es bestimmte Längenunterschiede $S-V$, welche nicht wahrgenommen werden.
3. Bei einem Standardreiz von $S=10$ cm werden kleine Längenunterschiede $S-V$ nicht wahrgenommen.
4. Bei einem Standardreiz von 10 cm Länge werden Längenunterschiede von weniger als 3 bis 4 mm nicht wahrgenommen.

Unter einer strengen Theorienprüfung versteht man, dass möglichst ernsthaft versucht werden soll, die Theorie zu falsifizieren bzw. ihren Bewährungsgrad zu senken.

Strenge Theorienprüfung

So wird z.B. die Geltung des Weber'schen Gesetzes für alle Sinnesmodalitäten und für alle möglichen Standardreize behauptet. Bisher haben wir es aber nur für den Längenvergleich und auch dabei nur für Standardreize von 10, 20, 30 und 40 cm überprüft.

Überprüfen wir das Weber'sche Gesetz nun erneut anhand derselben Standardreize, für die es sich schon einmal bewährt hat, so ist die Prüfung weniger streng, als wenn wir stattdessen sehr kleine (z.B. 1 bis 2 cm) und/oder sehr große Standardreize (z.B. 4 bis 5 m) verwenden. Es könnte ja sein, dass das Weber'sche Gesetz nur in einem mittleren Stimulusbereich (z.B. 10 bis 40 cm) gilt.

Selbst damit bleibt die Überprüfung des Weber'schen Gesetzes jedoch weniger streng, als wenn wir seine Geltung noch an weiteren Sinnesmodalitäten überprüfen, z.B. Gewichtswahrnehmung, Lautstärke usw.

Wissenschaft als sozialer Prozess

Methodische Norm und soziale Wirklichkeit

Die Forderung nach strenger Hypothesenprüfung ist wissenschaftspsychologisch und wissenschaftssoziologisch nicht ganz unproblematisch. So hat u.a. Holzkamp (1972) darauf hingewiesen, dass sich der Forscher, der die Hypothese aufgestellt hat, oft (unbewusst) dagegen sträubt, die Hypothese wirklich streng zu prüfen und so die "Vernichtung" (Falsifikation) seiner Hypothese zu riskieren. Es ist nur zu verlockend, das Experiment (bewusst oder unbewusst) so anzulegen, dass der vermutete Zusammenhang demonstriert wird.

Huber (1987, 60) versucht dies zu entkräften, indem er darauf hinweist, dass Wissenschaft auch ein sozialer Prozess ist: "Wenn eine Hypothese nur einigermaßen interessant ist, ruft sie auch andere Wissenschaftler auf den Plan. Selbst wenn der Vater oder die Mutter einer Hypothese eine ‚Beißhemmung‘ gegenüber dem eigenen Geisteskind haben sollte, andere Forscher haben diese Hemmung nicht. Immerhin kann man sich in der Wissenschaft durchaus einen Namen machen, wenn man eine bisher akzeptierte Hypothese falsifiziert. So wird durch den sozialen Kontext, in den die Hypothesenprüfung eingebettet ist, die strenge Prüfung gefördert".

Diese Argumentation mag beruhigen. Da sie jedoch auf einem methodischen Fehler beruht, geht sie leider an der Sache vor-

bei. Denn Huber weist zwar mit Recht darauf hin, dass Wissenschaft auch ein sozialer Prozess ist, stellt dann aber nicht die soziale Wirklichkeit der Scientific Community in Rechnung, sondern ein Idealbild von Wissenschaft als einer Institution, die völlig wertfrei nur der Wahrheit verpflichtet ist und in der Gleiche mit Gleichen in einem herrschaftsfreien Dialog stehen. Tatsächlich ist die Institution ‚Wissenschaft‘ (entgegen ihrem Erkenntnisideal) jedoch durchaus hierarchisch organisiert. Und wer als Wissenschaftler Karriere machen will, wird sich oft hüten, heilige Kühe zu schlachten usw.

Als methodische Norm bleibt die Forderung nach strenger Hypothesenprüfung davon unberührt. Gerade deshalb muss man sich aber der psychologischen und soziologischen Fallstricke bewusst sein, die ihrer Verwirklichung entgegenstehen. Wissenschaftler sind auch nur Menschen, die über diese Fallstricke stolpern können; vermeiden können sie dies am ehesten, wenn sie sich der Gefahren bewusst sind. Wenn man vor diesen Gefahren die Augen verschließt und sie beiseite schiebt, kann man mit Sicherheit damit rechnen, dass man ihnen früher oder später erliegt.

Unabdingbare Voraussetzung (strenger) Theorienprüfung ist die eindeutige Klärung der in der Theorie und der daraus abgeleiteten Sachhypothese vorkommenden Termini. Denn ohne explizit geklärte Terminologie wissen wir weder, was die Theorie aussagt, noch wissen wir, woran sie scheitern kann. Dass man es in der Psychologie damit lange Zeit nicht so ernst nahm, ist bis heute eine der verbreitetsten Fehlerquellen psychologischer Forschung.

Klärung der Terminologie als erste Voraussetzung strenger Theorienprüfung

Das schon von Wittgenstein (1953) kritisierte Nebeneinander von empirischen Methoden und Begriffsverwirrung hat in einigen Teilgebieten der Psychologie über lange Zeit zu einem generellen Theoriedefizit geführt, das sich z.B. in der von Holzkamp (1994) als Variablenpsychologie beschimpften Aneinanderreihung von zusammenhanglosen Einzelhypothesen der Form UV → AV bemerkbar machte. So unterschied man z.B. in der Aggressionsforschung eine Frustrations-Aggressions-Hypothese, eine Katharsis-Hypothese, eine Aggressions-Aggressions-Hypothese, eine Aggressions-Attributions-Hypothese, eine Bedrohungs-Aggressions-Hypothese usw., ohne sie in eine umfassende Theorie der

Aggression zu integrieren. Gleichzeitig schienen die empirischen Befunde teilweise widersprüchlich, weil jeder Autor die Aggression in seinen Experimenten anders operationalisierte.

Nicht weniger schmerzlich bemerkbar macht sich das teilweise immer noch bestehende Theoriedefizit darin, dass die Voraussetzungen, an welche die experimentelle Prüfung einer Theorie gebunden ist, oft keiner näheren Untersuchung unterzogen werden. Auch dies ist eine mehr oder minder direkte Folge der Vernachlässigung terminologischer Fragen. Denn ohne explizit geklärte Terminologie können weder Theorien auf ihre Widerspruchsfreiheit und/oder Kritisierbarkeit hin überprüft werden, noch kann die Kontingenz der im Experiment überprüften Sachhypothese sichergestellt werden, und selbst die Operationalisierbarkeit der theoretischen Begriffe ist nicht gegeben. Was bleibt, sind dann nur noch Ad-hoc-Operationalisierungen und mehr oder minder *plausibel* erscheinende Zusammenhänge zwischen unabhängigen und abhängigen Variablen, die durch diese Operationalisierungen – und nur durch sie – definiert sind; kurzum: eine Karikatur von Wissenschaft, die noch nicht einmal sicher sein kann, dass sie nicht bloße Pseudoempirie betreibt.

2.5.2.1 Widerspruchsfreiheit

Definition

Widerspruchsfreiheit bedeutet, dass der gesamte Komplex aus Theorie, Sachhypothese und Zusatzannahmen keine logischen oder sachlogischen Widersprüche enthalten darf. Formal:

Es darf nicht gleichzeitig A und $\neg A$ behauptet werden.

Die Forderung nach Widerspruchsfreiheit hat einen guten Grund, denn aus widerspruchshaltigen Theorien könnte man aufgrund des logisch wahren Satzes

$$[A \wedge \neg A] \rightarrow B$$

jede beliebige Schlussfolgerung ableiten, egal, ob B wahr ist oder nicht.

Auf den ersten Blick scheint das gleichzeitige Behaupten einer Aussage und ihrer Negation leicht erkennbar und die Forderung nach Widerspruchsfreiheit daher eher trivial. Tatsächlich ist dies auch der Fall, sobald die widersprüchlichen Aussagen einander

direkt gegenübergestellt werden. In der Praxis ist das Erkennen von Widersprüchen jedoch weit schwieriger. Denn Theorien bestehen in der Regel nicht nur aus einer Handvoll von Hypothesen, sondern sind relativ komplexe Systeme von Aussagen, wobei manche ihrer Behauptungen gar nicht explizit formuliert sind, sondern lediglich von anderen Aussagen impliziert werden.

Das Erkennen von Widersprüchen ist schon deutlich erschwert, wenn die widersprüchlichen Aussagen zwar explizit formuliert, aber räumlich voneinander getrennt sind, wie im folgenden Beispiel eines Zeitungsberichtes über alliierte Kriegsgefangene während des Golfkrieges (*The Times*, January 22, 1991: "Human shield ploy outrages comrades of missing airmen"; zit. n. Reimann, 2001).

Beispiel

- Im ersten Absatz des Berichtes wird verkündet: Die Kriegsgefangenen sind in Gefahr (A) ("Iraq threatens to use prisoners of war as human shields").
- Nachdem dies ausführlich als Verstoß gegen die Genfer Konvention und als erneuter Beweis für das Verbrechen Saddam Husseins ausgeschlachtet wurde, wird in Absatz 13/14 eine frühere britische Geisel als Zeuge dafür bemüht, dass die Kriegsgefangenen nichts zu befürchten haben ($\neg A$) ("had nothing to fear from the Iraqis").
- Daraus wird schließlich die Schlussfolgerung gezogen, dass die Alliierten ihre Bombenangriffe auf Bagdad unbedingt fortsetzen müssen (B). Genau so gut hätte man aber auch die umgekehrte Schlussfolgerung ziehen können ($\neg B$).

Das Erkennen des Widerspruches *threat* versus *nothing to fear* in unserem Beispiel wird auch dadurch erschwert, dass in den einander widersprechenden Aussagen kontradiktorische Prädikato- ren von unterschiedlicher Lautgestalt verwendet werden. Ob zwischen *threat* und *nothing to fear* wirklich ein Widerspruch besteht, ob die Aussagen wirklich kontradiktorisch sind, darüber lässt sich trefflich streiten, solange wir über keine geklärte Terminologie verfügen.

Ohne geklärte Terminologie kann nur die formale Widerspruchsfreiheit einer Theorie gewährleistet werden, und das ist nicht ausreichend, wie sich anhand eines Beispiels leicht zeigen lässt.

Formale
Widerspruchsfreiheit

Zu diesem Zweck betrachten wir die einfache Theorie

$$\forall x. \{[A(x) \wedge B(x)] \rightarrow C(x)\}.$$

Aus der Wahrheitstafel für $A(x) \wedge B(x)$ sehen wir, dass die Prämissen der Theorie formal widerspruchsfrei sind. Der darin dargestellte Sachverhalt $A(x) \wedge B(x)$ ist (formal) logisch möglich.

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

Sachlogische
Widerspruchsfreiheit

Um zu entscheiden, ob die Theorie tatsächlich widerspruchsfrei ist, müssen wir auch ihre sachlogische Widerspruchsfreiheit untersuchen. Dazu müssen wir wissen, was die Aussagen A und B bedeuten. Zum Beispiel bedeuten

$A(x) \Leftrightarrow$ "x ist ein Maluma"

und

$B(x) \Leftrightarrow$ "x ist eine Takete".

Damit sind wir jedoch noch nicht weiter gekommen, weil wir die Bedeutung unserer Termini ‚Maluma‘ und ‚Takete‘ noch nicht explizit vereinbart haben. In der Psychologie merken wir so etwas mitunter nicht gleich, weil wir oft Worte aus der Alltagssprache übernehmen und uns deren Lautgestalt bereits vertraut ist. In der Alltagssprache haben diese Worte zwar oft eine ähnliche, aber meist keine klar definierte Bedeutung. Wir verwenden die Alltagssprache, ohne uns explizit an Regeln der Wortverwendung zu halten. Die Regeln der Alltagssprache sind allenfalls implizit und meist unscharf.

Anhand der in Abb. 2.5.2 dargestellten Gegenstände können wir die Termini ‚Maluma‘ und ‚Takete‘ jedoch exemplarisch einführen und ihre Verwendungsweise durch Angabe von Prädikatenregeln weiter präzisieren.

Wir vereinbaren z.B. die Regeln "Malumas sind dick und rund" und "Taketen sind schmal und spitz". Die Worte ‚dick‘, ‚rund‘, ‚schmal‘ und ‚spitz‘ setzen wir dabei als bereits empirisch eingeübt voraus. Sollte sich ihre Verwendung als nicht hinreichend eindeutig erweisen, können wir ihre Einführung explizit wiederholen.

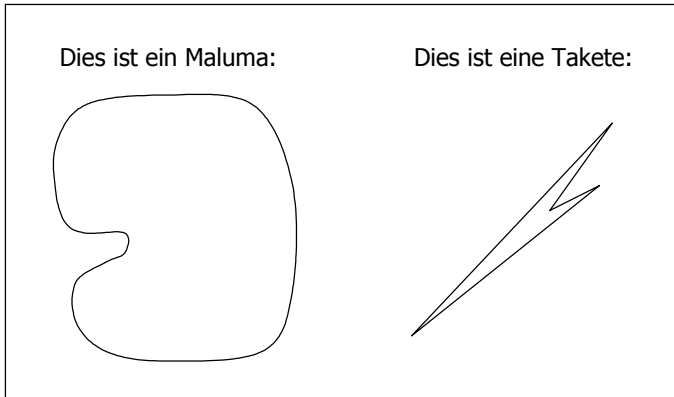


Abb. 2.5.2: Maluma und Takete

Vor allem aber formulieren wir die folgende Regel, wonach die Termini ‚Maluma‘ und ‚Takete‘ in einem konträren Gegensatz zueinander stehen:

$$A(x) \Rightarrow \neg B(x) \quad \text{"Malumas sind keine Taketen"} \\ \text{(und umgekehrt)}$$

Dann ist wegen der Prädikatorenregel $A(x) \Rightarrow \neg B(x)$ die Aussage $A \wedge B$ in sich widersprüchlich. Die Kombination $A \wedge B$ scheidet als sachlogisch unmögliches Ereignis aus unserer Wahrheitstafel aus, und die Aussage $A \wedge B$ ist stets falsch.

A	B	$A \wedge B$
w	f	f
f	w	f
f	f	f

Aus falschen Prämissen aber kann wieder jede beliebige Schlussfolgerung gezogen werden.

2.5.2.2 Kritizierbarkeit

Eine Theorie ist kritizierbar, wenn es mögliche Ergebnisse ihrer Prüfung gibt, welche die Theorie falsifizieren oder ihren Bestätigungsgrad reduzieren.

Definition

Nicht kritisierbar sind solche Theorien, die unter allen denkbaren Umständen wahr sind. Eine derartige Theorie ist immun gegen Prüfung. Sie ist zwar stets wahr, erklärt aber nichts.

Beispiele

Kritisierbar ist z.B. die entwicklungspsychologische Theorie, wonach Kinder unter sieben Jahren Kopffüßler zeichnen. Wenn sich Kinder finden lassen, die bereits unter sieben Jahren keine Kopffüßler mehr zeichnen, so ist die Theorie falsifiziert.

Als anschauliches Beispiel nicht kritisierbarer Theorien führt Huber (1987, 54) die mittelalterlichen Hexenprozesse an: Die Anschuldigung (Theorie) lautete z.B., Frau M.S. stehe mit einer teuflischen Macht im Bunde. Zur Überprüfung der Anschuldigung wurde Frau M.S. verhört. Gestand sie, dass sie mit dem Teufel im Bunde sei, war die Anschuldigung klarerweise bestätigt. Gestand sie nicht, wurde sie gefoltert. Gestand sie im Laufe der Folter, war die Anschuldigung ebenfalls bestätigt. Gestand sie auch unter der härtesten Folter nicht, so war das nur dadurch zu erklären, dass ihr der Teufel während der Folter beigestanden war. Auch in diesem Fall war das Bündnis mit dem Teufel erwiesen und die Anschuldigung damit bestätigt.

Kritisierbarkeit und logische Form

Bestimmte Theorien sind bereits aufgrund ihrer logischen Form nicht kritisierbar. Dazu gehören insbesondere:

- widerspruchshaltige Theorien (siehe oben),
- Theorien, deren Conclusio tautologisch ist, und
- Theorien, die auf einem Zirkelschluss beruhen.

Widerspruchshaltige Theorien

Widerspruchshaltige Theorien sind nicht kritisierbar, weil sich aus ihnen keine Sachhypothese (B) herleiten lässt, welche die Theorie falsifizieren könnte, denn es ist stets beides wahr:

sowohl $[A \wedge \neg A] \rightarrow B$

als auch $[A \wedge \neg A] \rightarrow \neg B$.

Tautologische Conclusio

Theorien, deren Conclusio tautologisch ist, sind stets wahr, weil die Subjunktion $A \rightarrow B$ stets wahr ist, wenn der Nachsatz (B) wahr ist – und zwar unabhängig davon, ob die Prämissen (A) wahr sind oder nicht. Z.B. die alte Bauernregel

$A \rightarrow [C \vee \neg C]$.

"Kräht der Hahn auf dem Mist (A), ändert sich das Wetter (C), oder es bleibt, wie es ist ($\neg C$)".

Theorien, die auf einem Zirkelschluss der Form $A \rightarrow A$ beruhen, sind nicht kritisierbar, weil die Subjunktion $A \rightarrow A$ stets wahr ist – und zwar unabhängig davon, ob A wahr ist oder nicht. Auf einem solchen Zirkelschluss beruht z.B. die vermögenspsychologische Verdoppelung der Realität. Zirkelschluss

Auch dabei ist es wieder so, dass ohne geklärte Terminologie nur die formale Kritisierbarkeit gewährleistet werden kann, was jedoch nicht ausreichend ist. Formale und sachlogische Kritisierbarkeit

Zum Beispiel betrachten wir die formal kritisierbare Theorie

$$\forall x. \{A(x) \rightarrow \neg B(x)\}.$$

Formal kritisierbar bedeutet dabei so viel wie von ihrer logischen Form her, d.h. ohne dass wir die Bedeutung der Aussagen A und B beachten:

A	B	$\neg B$	$A \rightarrow \neg B$
w	w	f	f
w	f	w	w
f	w	f	w
f	f	w	w

Wie wir der Wahrheitstafel entnehmen, kann das Ereignis $A \wedge B$ die Theorie falsifizieren. Damit die Theorie tatsächlich kritisierbar ist, muss dieses Ergebnis auch sachlogisch möglich sein. Um dies entscheiden zu können, müssen wir wissen, was A und B bedeuten. Zum Beispiel bedeuten

$$A(x) \Leftrightarrow \text{"x ist ein Maluma"}$$

und

$$B(x) \Leftrightarrow \text{"x ist eine Takete"}.$$

Wegen der Regel $A(x) \Rightarrow \neg B(x)$ ("Malumas sind keine Taketen") scheidet aus unserer Wahrheitstafel aber die Kombination $A \wedge B$ als sachlogisch unmögliches Ereignis aus, und die Theorie ist nicht mehr kritisierbar.

A	B	$\neg B$	$A \rightarrow \neg B$
w	f	w	w
f	w	f	w
f	f	w	w

Verdoppelung der
Realität

Die naive Triebtheorie der Aggression

Als Beispiel einer Realitätsverdoppelung untersuchen wir die naive Triebtheorie der Aggression, wie sie uns im Alltagsdiskurs oft begegnet: Das beobachtete Verhalten (Aggression) wird zum Aggressionstrieb vergegenständlicht und dann durch seine Vergegenständlichung erklärt.

Grundlage für die Annahme eines Aggressionstriebes ist das Auftreten von Aggression: Menschen haben einen Aggressionstrieb, weil sie aggressiv handeln. Die Beobachtung von Aggression erlaubt es also, von einem Aggressionstrieb zu sprechen. Formal:

$$A \Rightarrow B.$$

Als Ursache der Aggression (A) wird dann der Aggressionstrieb (B) angenommen: Menschen handeln aggressiv, weil sie einen Aggressionstrieb haben. Formal:

$$B \rightarrow A.$$

Die Zirkularität der naiven Triebtheorie der Aggression ergibt sich schließlich aus der Zusammensetzung der beiden Teilaussagen:

- Aufgrund der Regel $A \Rightarrow B$ sprechen wir von einem Aggressionstrieb, wenn wir Aggression beobachten, und
- gemäß der "Theorie" $B \rightarrow A$ "erklären" wir diese Aggression durch den Aggressionstrieb.

Fazit:

$$\begin{array}{l} A \Rightarrow B \\ B \rightarrow A \\ \hline A \rightarrow A \end{array}$$

Menschen handeln aggressiv, weil sie aggressiv handeln.

Konrad Lorenz (1963), der Begründer der Aggressionstriebeorie, war freilich nicht so naiv, sondern machte die Annahme

eines Aggressionstriebes an bestimmten Gesetzmäßigkeiten des aggressiven Verhaltens fest:

- autonome Triebproduktion (Staubarkeit der Aggression), und
- Katharsis (Auslebbarkeit der Aggression).

Beides hat sich jedoch empirisch nicht bestätigt. Die Triebtheorie der Aggression wird daher selbst in der Biologie heute nicht mehr vertreten.

2.5.2.3 Exkurs: Die Triebtheorie der Aggression nach Konrad Lorenz

Laut Konrad Lorenz (1963) beruht die Aggression auf einem endogen produzierten Trieb, der – wie Hunger oder Sexualität – in periodischen Abständen sein Recht verlangt und allenfalls sublimiert, nicht aber durch Erziehungseinflüsse aus der Welt geschafft werden kann.

Staubarkeit der Aggression

Dabei zählt Lorenz die Aggression zum Instinktverhalten, dessen Auslösung die Überwindung einer Hemmung voraussetzt, wozu ein angeborener Auslösemechanismus (AAM) dient, der auf Schlüsselreize anspricht. Der Schlüsselreiz für das Angriffsverhalten des männlichen Stichlings besteht z.B. in einem roten Fleck am Bauch eines anderen männlichen Stichlings.

Diese Reizkonstellation ist äußerst spezifisch: Das Angriffsverhalten bleibt aus, wenn sich der rote Fleck etwa am Rücken des Rivalen befindet. Aber auch in Gegenwart eines Schlüsselreizes tritt das Instinktverhalten nicht immer auf: Tiere kämpfen in Reaktion auf aggressionsauslösende Signale nicht permanent, sondern nur, wenn sie in Stimmung sind, wofür Lorenz eine permanente Produktion aggressionspezifischer Triebenergie durch den Organismus annimmt.

Durch die Ausführung von Aggressionsverhalten wird dagegen aggressive Triebenergie verbraucht, und zwar desto mehr, je heftiger die Aggression ist. Fehlt es an aggressionsauslösenden Schlüsselreizen, so staut sich die aggressive Energie auf.

Je mehr aggressive Triebenergie aufgestaut wurde, desto weniger wirksame Schlüsselreize sind erforderlich, um Aggressionsverhalten auszulösen: Das Aggressionsverhalten tritt auf, wenn

die Summe aus der Wirksamkeit des Schlüsselreizes und dem Ausmaß der aufgestauten Triebenergie gegenüber der Hemmung überwiegen.

Im Extremfall einer sehr langen Stauung kommt es zu einer eruptiven Entladung der aggressiven Triebenergie, die verheerende Folgen haben kann.

Isoliert man z.B. einen Buntbarsch zusammen mit seinem Weibchen und nimmt ihm dadurch – so Lorenz – jeden Auslöserreiz für Aggression, so greift das Männchen bald auch sein eigenes Weibchen an, bis hin zu dessen Tod (Rasa, 1969).

Auslebbarkeit der Aggression

Da die Produktion der aggressiven Triebenergie nicht gedrosselt werden kann, muss die Aggression ausgelebt werden, wofür Lorenz u.a. aggressive Kampfsportarten vorschlägt. Durch kognitive Kontrolle der Aggression wird die Wahrscheinlichkeit für immer größere Katastrophen nur erhöht. Ihre Wirkung ist, "als wollte man dem Ansteigen des Dampfdruckes in einem dauernd geheizten Kessel dadurch begegnen, dass man am Sicherheitsventil die Verschlussfeder fester schraubt" (Lorenz, 1974, S.247).

Isolationsexperimente an Fischen

Nach dem heutigen Stand der verhaltensbiologischen Aggressionsforschung jedoch kann weder die Staubarkeit noch die Auslebbarkeit der Aggression als erwiesen gelten.

Die soziale Isolation in verschiedenen Fischgruppen und in verschiedenen Phasen der Ontogenese führt nach Franck (1985) zu unterschiedlichen Veränderungen der Aggressionsbereitschaft:

- Soziale Isolation erwachsener Männchen bewirkte bei manchen Fischarten eine erhöhte, bei anderen eine herabgesetzte Aggressionsbereitschaft.
- Lediglich die Aggressionsbereitschaft junger, noch nicht geschlechtsreifer Fische stieg bei allen bisher untersuchten Arten nach sozialer Isolation an.

Gleichermaßen lässt sich nach Franck (1985) die Interpretation derartiger Isolationsexperimente im Sinne einer Stauung aggressiver Energie nicht aufrechterhalten. Denn: Soziale Isolation verhindert nicht nur aggressives Verhalten, sondern jegliche sozialen Verhaltensweisen. Die Umwelt eines sozial isolierten Tieres ist insgesamt viel reizärmer als im sozialen Verband. Die Isolationseffekte lassen sich daher nicht auf einen einzigen Faktor in der sozialen Umwelt eines Tieres zurückführen.

Hinzu kommt, dass die Aggression der Jungfische eine ganz andere Funktion erfüllt als die Aggression erwachsener Männchen. Während die Aggression erwachsener Männchen vor allem der Abgrenzung eines Fortpflanzungsterritoriums dient und von der Konzentration der männlichen Sexualhormone abhängig ist, dient die Aggression der Jungfische vermutlich dazu, dem heranwachsenden Tier einen möglichst guten Zugang zur Nahrung zu sichern und dürfte somit unabhängig von den männlichen Sexualhormonen auftreten.

Die von Lorenz (1963) als Beleg für die behauptete Staubarkeit der Aggression herangezogenen Verhaltensbeobachtungen an Buntbarschen können deshalb als nicht schlusskräftig zurückgewiesen werden.

Auch die Vorstellung, dass größere Katastrophen (wie z.B. Kriege) durch Ausleben der Aggression (z.B. durch aggressive Kampfsportarten) vermieden werden könnten, wurde inzwischen durch ethnologische Studien (Sipes, 1973) empirisch widerlegt.

Anthropologische Studien zur Auslebbarkeit der Aggression

Je zehn relativ kriegerische und relativ friedliche Gesellschaften wurden daraufhin untersucht, ob sie Kampfsportarten ausüben. Als Basis für die Auswahl dieser Gesellschaften diente die internationale Kriegsstudie von Otterbein (1968), die zwischen drei Formen des Krieges unterscheidet: interne Kriege, externe Angriffskriege und externe Verteidigungskriege. Jede dieser Kriegsorten hatte Otterbein daraufhin beurteilt, ob sie in einer repräsentativen Stichprobe von 50 auf der Grundlage des ethnographischen Atlas' von Murdock (1962-64) ausgewählten Gesellschaften *selten oder nie*, *häufig* oder *kontinuierlich* auftreten.

- Als kriegerisch beurteilte Sipes solche Gesellschaften, die häufig oder kontinuierlich Angriffskriege führen.
- Als friedlich beurteilte er solche Gesellschaften, in welchen alle drei Kriegsformen selten oder nie vorkommen.

Als Kampfsportarten wurden solche Sportarten definiert, in denen zwei Opponenten (Individuen oder Mannschaften) gegeneinander antreten und bei denen es zu direktem oder indirektem (durch echte oder simulierte Waffen vermitteltem) Körperkontakt mit dem Ziel einer tatsächlichen oder symbolischen Verletzung des Gegners und/oder der Eroberung von Territorium kommt, oder in denen (unter Einsatz tatsächlicher oder simulierter Waf-

fen gegen echte oder simulierte Menschen) kriegsähnliche Aktivitäten ausgeführt werden.

Die Ergebnisse der Untersuchung zeigten, dass neun der zehn kriegerischen Gesellschaften, aber nur zwei der zehn friedlichen Gesellschaften Kampfsportarten ausübten. Fazit: Kriege werden durch die Ausübung von Kampfsportarten nicht vermieden, sondern im Gegenteil sind solche Gesellschaften, welche Kampfsportarten ausüben, insgesamt kriegerischer.

Schlussfolgerungen

Wenngleich das von Konrad Lorenz entworfene "Dampfkesselmodell" des Aggressionstriebes von vielen seiner Kritiker belächelt wurde und die Triebtheorie der Aggression schließlich an der Erfahrung gescheitert ist, kann die Leistung, welche Konrad Lorenz mit der Entwicklung seiner Theorie erbracht hat, nicht genug gewürdigt werden: Erst durch die Übersetzung des – zunächst auf einer Verdoppelung der Realität beruhenden – Konzepts eines Aggressionstriebes in eine – trotz allem – relativ komplexe, widerspruchsfreie, nicht-tautologische und nicht zirkuläre Theorie wurde die Vorstellung von einem Aggressionstrieb kritisierbar und die Grundlage dafür gelegt, dass die Theorie an der Erfahrung scheitern konnte.

Damit die Theorie an der Erfahrung überprüft werden konnte, waren noch andere Voraussetzungen zu erfüllen:

1. die Ableitbarkeit konkreter Sachhypothesen,
2. die Kontingenz der abgeleiteten Sachhypothesen sowie
3. die Operationalisierbarkeit der Sachhypothesen.

2.5.2.4 Kontingenz

Definition

Eine Sachhypothese ist kontingent, wenn sowohl das Bestehen des darin behaupteten Zusammenhanges als auch sein Nichtbestehen (sach)logisch möglich ist.

Wenn es sich bei der Sachhypothese um einen strukturellen Zusammenhang handelt, kann dieser nicht an der Erfahrung scheitern. Die Sachhypothese ist in diesem Fall nicht falsifizierbar und somit auch nicht geeignet, um die Geltung der Theorie zu prüfen, aus welcher sie abgeleitet wurde. Man muss daher sicherstellen, dass zwischen UV und AV kein struktureller Zusammenhang besteht.

Als Beispiel für Verletzung der Kontingenz führt Holzkamp (1986) Beispiele die folgende Hypothese nach Crott (1979, 169) an:

"Bei lohnenden Gewinnaussichten neigen die Vpn ... dazu, Lösungen anzustreben, die ihnen einen möglichst hohen Gewinn einbringen" (1979, 169),

und schlägt zwei einfache Testfragen vor, mittels derer man den Verdacht eines strukturellen Zusammenhanges zwischen UV (lohnende Gewinnaussicht) und AV (Anstreben einer Lösung, die möglichst hohen Gewinn einbringt) erhärten kann:

1. Prüfe, ob das Einfügen des Wörtchens ‚vernünftigerweise‘ sinnvoll möglich ist:

"Bei lohnenden Gewinnaussichten neigen die Vpn vernünftigerweise dazu, Lösungen anzustreben, die ihnen einen möglichst hohen Gewinn einbringen".

Ist dies der Fall, so besteht zwischen UV und AV möglicherweise kein empirischer (Bedingungs-)Zusammenhang, sondern ein (sachlogischer) Begründungszusammenhang.

2. Prüfe, ob das Einfügen des Wörtchens ‚nicht‘ sinnvoll möglich ist, d.h. eine empirisch zur Disposition stehende Alternative ergibt:

"Bei lohnenden Gewinnaussichten neigen die Vpn nicht dazu, Lösungen anzustreben, die ihnen einen möglichst hohen Gewinn einbringen".

Erscheint die Negation des Zusammenhanges als widersinnig, so ist der Zusammenhang zwischen UV und AV vermutlich strukturell.

In der Psychologie ist die Kontingenz der Sachhypothese vor allem dann nicht unproblematisch, wenn auf Seiten der unabhängigen und/oder der abhängigen Variable subjektseitig definierte Sinngehalte (wie z.B. die Wahrnehmung einer Situation, ihre Deutung als Konfliktsituation, die Interpretation des Konfliktes als Konkurrenzsituation [win-lose-Modell], die Intention, sich durchzusetzen; die Einschätzung bestimmter Verhaltensweisen als geeignetes Mittel usw.) ins Spiel kommen. Wie wir im Kontext der intentionalen Erklärung sehen werden, sind Sinngehalte sowohl untereinander als auch mit objektseitig definierten Sachverhalten durch Begründungszusammenhänge verbunden. Die Gefahr eines strukturellen Zusammenhanges zwischen UV und AV ist in diesem Fall daher stets gegeben (vgl. Tab. 2.5.6). Sinngehalte

Formale und sachlogische Kontingenz

Ohne geklärte Terminologie kann auch hier nur die formale Kontingenz gewährleistet werden. Das ist jedoch nicht ausreichend.

Zum Beispiel betrachten wir die formal kontingente Sachhypothese

$$\forall x. \{A(x) \rightarrow \neg B(x)\}.$$

Sowohl das Ereignis $A \wedge \neg B$ als auch das Ereignis $A \wedge B$ ist (formal) logisch möglich, und die Subjunktion $A \rightarrow \neg B$ kann (formal) logisch bestehen oder auch nicht:

A	B	$\neg B$	$A \rightarrow \neg B$
w	w	f	f
w	f	w	w
f	w	f	w
f	f	w	w

Tab. 2.5.6: Unabhängige Variable, abhängige Variable und Kontingenz der Zusammenhangshypothese

Unabhängige Variable	Abhängige Variable	Kontingenz
Physikalischer Stimulus, z.B. Lichtreiz	Physiologische Reaktion, z.B. elektrische Gehirnaktivität	Unproblematisch
Physikalischer Stimulus, z.B. Lichtreiz	Beobachtbares Verhalten, z.B. Lidschlagreflex	Unproblematisch
Physikalischer Stimulus, z.B. Unterschied zwischen zwei Lichtreizen	Sinnhaft gedeutetes Verhalten, z.B. Unterschiedswahrnehmung	Unproblematisch
Physikalischer Stimulus, z.B. Bewegung zweier Lichtpunkte	Sinngehalt, z.B. deren Deutung als soziale Interaktion	Unproblematisch
Physiologischer Vorgang	Physiologischer Vorgang	Nicht Gegenstand der Psychologie
Physiologischer Vorgang, z.B. Alpha-Wellen	Beobachtbares Verhalten, z.B. Tiefschlaf	Unproblematisch
Physiologischer Vorgang, z.B. Aktivität einer best. Hirnregion	Sinnhaft gedeutetes Verhalten, z.B. Flucht	Unproblematisch
Physiologischer Vorgang, z.B. Vergiftung	Sinngehalt, z.B. Halluzination	Unproblematisch

Unabhängige Variable	Abhängige Variable	Kontingenz
Beobachtbares Verhalten, z.B. Schreckreaktion	Physiologisches Korrelat, z.B. Adrenalinausschüttung	Unproblematisch
Beobachtbares Verhalten, z.B. Alkoholgenuss	Beobachtbares Verhalten, z.B. Gleichgewichtsstörung	Unproblematisch
Beobachtbares Verhalten, z.B. (Weg-)Laufen	Sinnhaft gedeutetes Verhalten, z.B. Flucht	ACHTUNG
Beobachtbares Verhalten, z.B. (Weg-)Laufen	Sinngehalt, z.B. 'Bedrohung'	ACHTUNG
Sinnhaft gedeuteter Stimulus, z.B. Frustration	Physiologische Reaktion, z.B. Blutdrucksteigerung	Unproblematisch
Sinnhaft gedeuteter Stimulus, z.B. Frustration	Beobachtbares Verhalten, z.B. Weinen	Unproblematisch
Sinnhaft gedeuteter Stimulus, z.B. Frustration	Sinnhaft gedeutetes Verhalten, z.B. Aggression	ACHTUNG
Sinnhaft gedeuteter Stimulus, z.B. Frustration	Sinngehalt, z.B. 'Konflikt'	ACHTUNG
Sinnhaft gedeutetes Verhalten, z.B. Aggression	Physiologisches Korrelat, z.B. Blutdrucksteigerung	Unproblematisch
Sinnhaft gedeutetes Verhalten, z.B. Flucht	Beobachtbares Verhalten, z.B. (Weg-)Laufen	ACHTUNG
Sinnhaft gedeutetes Verhalten, z.B. Aggression	Sinnhaft gedeutetes Verhalten, z.B. Durchsetzung	ACHTUNG
Sinnhaft gedeutetes Verhalten, z.B. Flucht	Sinngehalt, z.B. 'Bedrohung'	ACHTUNG
Sinngehalt, z.B. 'Bedrohung'	Physiologisches Korrelat, z.B. Adrenalinausschüttung	Unproblematisch
Sinngehalt, z.B. 'Bedrohung'	Beobachtbares Verhalten, z.B. (Weg-)Laufen	Unproblematisch
Sinngehalt, z.B. 'Bedrohung'	Sinnhaft gedeutetes Verhalten, z.B. Flucht	ACHTUNG
Sinngehalt, z.B. 'Bedrohung'	Sinngehalt, z.B. 'Konflikt'	ACHTUNG

Um zu entscheiden, ob die Sachhypothese $A \rightarrow \neg B$ tatsächlich kontingent ist, müssen wir aber wissen, ob das Ereignis $A \wedge B$ überhaupt sachlogisch möglich ist. Dafür müssen wir wieder wissen, was A und B bedeuten. Zum Beispiel bedeutet

$A(x) \Leftrightarrow$ "x ist ein Maluma"

und

$B(x) \Leftrightarrow$ "x ist eine Takete".

Dann ist wegen der Prädikatorenregel $A(x) \Rightarrow \neg B(x)$ ("Malumas sind keine Taketen") die Aussage $A \wedge B$ in sich widersprüchlich. Die Kombination $A \wedge B$ scheidet als sachlogisch unmögliches Ereignis aus unserer Wahrheitstafel aus, und die Hypothese $A \rightarrow \neg B$ ist (ohne dass wir dies durch Beobachtung zu bestätigen brauchen) stets wahr. D.h. die Hypothese ist nicht kontingent, und jeder Versuch ihrer empirischen Prüfung bloße Pseudo-Empirie.

A	B	$\neg B$	$A \rightarrow \neg B$
w	f	w	w
f	w	f	w
f	f	w	w

Die sachlogische
Begründung
struktureller
Gesetzmäßigkeiten

Mithilfe von Holzkamps Testfragen lässt sich die Kontingenz einer Hypothese begründet anzweifeln. Etwa im folgenden Beispiel

"Wenn jemand in einer win-lose-Situation nicht nachgibt, wird er versuchen, sich durchzusetzen",

bei dem das Einschieben des Wörtchens ‚nicht‘ keine empirisch zur Disposition stehende Alternative ergibt:

"Wenn jemand in einer win-lose-Situation nicht nachgibt, wird er nicht versuchen, sich durchzusetzen".

Der definitive Nachweis, dass es sich dabei um keine kontingente Hypothese handelt, welche es empirisch zu überprüfen gilt, sondern um eine strukturelle Gesetzmäßigkeit, die sachlogisch zu begründen ist, lässt sich aber nur erbringen, indem man eben diese Begründung vorlegt. Um dies zu leisten, werden wir ggf. unsere wissenschaftliche Terminologie genauer klären müssen: Was heißt ‚win-lose-Situation‘, was heißt ‚sich durchsetzen‘, was heißt ‚nachgeben‘?

2.5.2.5 Operationalisierbarkeit

<p>Eine (empirische) Hypothese ist operationalisierbar, wenn den darin vorkommenden Begriffen (Termini) beobachtbare Daten zugeordnet werden können.</p>	Definition
<p>Die Operationalisierung der Hypothese erweist sich als unproblematisch, wenn UV und AV objektseitig definierte und beobachtungssprachlich beschriebene Sachverhalte sind. Schwieriger ist es, wenn UV oder AV theoriesprachlich beschrieben werden, und/oder wenn sie subjektseitig definiert sind.</p>	Voraussetzungen
<p>Unabdingbare Voraussetzung ist es in beiden Fällen, dass die in der Hypothese vorkommenden Termini in ihrer Bedeutung eindeutig geklärt sind und dass die Operationalisierung der Termini ihrer Definition entspricht. Nur wenn die Bedeutung der Termini geklärt ist, kann man entscheiden, ob die im Experiment verwendeten Operationalisierungen auch dem entsprechen, wovon die Theorie handelt.</p>	
<p>So ist es z.B. ungenügend, Aggression als (versuchte) Durchsetzung eigener Handlungen und/oder Ziele gegen den Willen eines anderen zu definieren und die so definierte Aggression im Experiment dadurch zu operationalisieren, dass ein Kind seine Puppe schlägt.</p>	Ein Beispiel
<p>Das Schlagen der Puppe könnte auch z.B. Rache oder sinnlose Gewalt sein. Zu einer Operationalisierung von Aggression im konflikttheoretischen Sinne wird das Schlagen der Puppe erst, wenn es in einen Spielkontext eingebettet ist, in dem das Kind z.B. die Puppe schlägt, um sie zu erziehen, d.h. damit sie sich in Zukunft in gewünschter Weise verhält (z.B. nicht mehr einnässt).</p>	
<p>Beschreiben wir z.B. Pawlows Experimente mittels der objektseitig definierten beobachtungssprachlichen Prädikatoren ‚Glockenton‘, ‚Futter‘ und ‚Speichelfluss‘, so stellt sich das Problem der Operationalisierung überhaupt nicht. Die UV und AV können unmittelbar beobachtet bzw. gemessen werden.</p>	UV und AV beobachtungssprachlich und objektseitig definiert
<p>Schwieriger wird es dagegen, wenn wir von Pawlows konkreten Experimenten abstrahieren und den bedingten Reflex mittels der theoriesprachlichen Termini ‚neutraler Stimulus‘, ‚unkonditionierter Stimulus‘ und ‚unkonditionierte Reaktion‘ darstellen.</p>	UV und AV theoriesprachlich und objektseitig definiert

Außer der Beobachtung des Glockentons, des Futters und des Speichelflusses werden bestimmte Zusatzannahmen erforderlich, die empirisch zu prüfen sind: Dass der Glockenton ein NS ist, setzt voraus, dass er den Speichelfluss zunächst nicht auslöst. Dass das Futter ein US ist, setzt voraus, dass es den Speichelfluss regelmäßig auslöst. Dass der Speichelfluss eine UR auf das Futter ist, setzt voraus, dass er durch den Anblick des Futters ausgelöst wird.

AV beobachtungssprachlich, aber subjektseitig definiert

Obwohl es sich dabei um eine beobachtungssprachliche Variable handelt, kann auch die AV in den Experimenten zum Weber'schen Gesetz nicht ohne Zusatzannahmen operationalisiert werden.

Um uns auf die Aussagen der Probanden, der Vergleichsreiz sei "länger" bzw. "kürzer" als oder "gleich lang" wie der Standardreiz, als Indikatoren für die Wahrnehmung $w(S-V)$ verlassen zu können, müssen wir voraussetzen, (1) dass die Probanden die Worte ‚länger‘, ‚kürzer‘ und ‚gleich lang‘ richtig gebrauchen, (2) dass sie Standardreiz und Vergleichsreiz nicht verwechseln sowie (3) dass sie wahrheitsgemäß antworten, d.h. dass sie tatsächlich das beschreiben, was sie sehen.

Verzicht auf Selbstauskünfte

Der Rückgriff auf die Selbstauskünfte der Probanden stellt nur eine Möglichkeit der Operationalisierung subjektseitig definierter AVn dar. So könnte man zwecks Überprüfung des Weber'schen Gesetzes im Tierversuch die Wahrnehmung des Stimulusunterschieds z.B. wie folgt operationalisieren:

1. Auf den Standardreiz wird eine Reaktion R_1 konditioniert, z.B. Speichelfluss (Futter).
2. Auf den Vergleichsreiz wird eine andere Reaktion R_2 konditioniert, z.B. Fluchtreaktion (E-Schock).
3. Gelingt es, beide Reaktionen gleichzeitig an demselben Versuchstier zu konditionieren, so kann darauf geschlossen werden, dass das Versuchstier den Stimulusunterschied wahrnimmt $w(S-V)$.

Die dazu erforderlichen Zusatzannahmen sind: (1) S ist ein NS bezüglich R_1 . (2) V ist ein NS bezüglich R_2 . (3) R_1 kann auf S konditioniert werden ($w(S)$). (4) R_2 kann auf V konditioniert werden ($w(V)$). (5) R_1 und R_2 können gleichzeitig auf verschiedene Stimuli konditioniert werden.

Alle diese Zusatzannahmen sind empirisch zu prüfen, bevor das eigentliche Experiment zum Weber'schen Gesetz stattfinden kann.

Als Beispiel für theoriesprachlich und/oder subjektseitig definierte UVn kann die Frustrations-Aggressions-Hypothese (Dollard et al., 1939) dienen. Darin wird Frustration als Unterbrechung einer zielgerichteten Aktivität definiert, und es ergibt sich das Problem, dass Handlungsziele nicht beobachtbar sind, sondern eine intentionale Deutung des beobachteten Verhaltens beinhalten: Das Verhalten wird ausgeführt, um ein bestimmtes Ziel zu erreichen.

UV theoriesprachlich und/oder subjektseitig definiert

- Im Experiment würde man die Frustration z. B. so operationalisieren, dass man den Probanden eine Aufgabe stellt und ihnen im Falle der erfolgreichen Bewältigung der Aufgabe einen bestimmten Geldbetrag als Belohnung in Aussicht stellt.
- Tatsächlich wird die Versuchsanordnung aber so eingerichtet, dass die Probanden bei der Bearbeitung der Aufgabe gestört werden, sie deshalb nicht erfolgreich zu Ende führen können und auch keine Belohnung erhalten.

Das erscheint zwar durchaus plausibel. Welche Zusatzannahmen dazu jedoch erforderlich sind, können wir mit den bisher erarbeiteten terminologischen Mitteln noch gar nicht darstellen.

Mit der Deutung einer Situation als Frustration gehen wir nämlich über das deduktiv-nomologische Erklärungsmodell hinaus und bedienen uns einer intentionalen Erklärung, so dass in der Frustrations-Aggressions-Hypothese – selbst bei Zugrundelegung des behavioristischen Aggressionsbegriffs – in die deduktiv-nomologische Erklärung

$$\forall x. x \in \text{Frustration} \rightarrow x \in \text{Aggression}$$

eine intentionale Erklärung

$$x \in \text{Frustration}$$

eingebettet ist. Welche methodologischen Konsequenzen das hat, werden wir später diskutieren; insbesondere die Frage, ob bzw. unter welchen Voraussetzungen sich auch im Tierversuch von zielgerichtetem Verhalten sprechen lässt.

Zusammenfassend sei hier nur so viel festgehalten:

Fazit

1. Operationalisierung setzt eine geklärte Terminologie voraus.
2. Ad-hoc-Operationalisierungen, die auf bloßer Plausibilität beruhen, reichen nicht aus.

3. Obwohl häufig als operationale Definitionen bezeichnet, sind Operationalisierungen kein Ersatz für die Klärung der Termini, sondern im Gegenteil ist die Klärung der Terminologie eine unverzichtbare Voraussetzung für angemessene Operationalisierungen.
4. Hypothesen, die theoriesprachliche Termini enthalten, können niemals isoliert geprüft werden, denn ihre empirische bzw. experimentelle Überprüfung setzt stets gewisse Zusatzannahmen voraus.

Kontrollfragen

- Wozu dient das Experiment im Kontext deduktiv-nomologischer Erklärungen?
- Worin besteht das Experiment? Geben Sie ein Beispiel!
- Welchem Element des deduktiv-nomologischen Erklärungsmodells entspricht im Experiment die UV? Welchem entspricht die AV?
- Warum können empirische Gesetzesaussagen niemals definitiv bewiesen werden?
- Was versteht man unter einer Hypothese?
- Warum bleiben empirische Gesetzesaussagen stets Hypothesen?
- Wodurch können sie falsifiziert werden, und welcher logischen Schlussform folgt diese Falsifikation?
- Zeigen Sie, dass diese Schlussform logisch wahr ist!
- Was bedeutet die Aussage, dass sich eine Hypothese bewährt hat?
- Wann gilt eine Hypothese als gut bewährt, wann als weniger gut bewährt?
- Was versteht man unter den Stufen einer Variable?
- Worin besteht der Unterschied zwischen einfaktoriellem und mehrfaktoriellem Experimenten?
- Woran ist abzulesen, ob sich die einem Experiment zugrundeliegenden Hypothesen bewährt haben oder nicht, und woraus ist abzulesen, welche der Hypothesen ggf. falsch sind?
- Was versteht man unter der Generalisierung der im Experiment gewonnenen Ergebnisse?
- Auf welcher Art von Schluss beruht die Generalisierung?
- Worin unterscheidet sich die Generalisierung von den Schlussformen des modus ponens und des modus tollens? Was folgt daraus?
- Worin besteht der Unterschied zwischen induktiven und deduktiven Schlüssen?
- Welche methodische Reihenfolge besteht bezüglich der Formulierung und Prüfung von Hypothesen?
- Was versteht man unter einer Störvariable?
- Was versteht man unter der Konfundierung von SV_n und UV_n , und was folgt daraus?
- Was versteht man unter der internen Validität eines Experimentes?
- Was versteht man unter der externen Validität eines Experimentes?

- Worin besteht der Unterschied zwischen Laborexperimenten, Feldexperimenten und Feldstudien? Welche Vor- bzw. Nachteile haben sie bezüglich der internen bzw. externen Validität der Untersuchungsergebnisse?
- Beschreiben Sie das Ferienlagerexperiment von Sherif!
- Erläutern Sie den Unterschied zwischen univariaten und multivariaten Experimenten!
- Erläutern Sie den Unterschied zwischen Versuchsplänen mit unabhängigen Stichproben und Versuchsplänen mit abhängigen Stichproben!
- Erläutern Sie die Termini ‚Versuchsbedingung‘ und ‚Kontrollbedingung‘!
- Erläutern Sie die Unterscheidung zwischen Versuchsgruppe und Kontrollgruppe!
- Beschreiben Sie Versuchsaufbau und (einige) Ergebnisse des Milgram-Experimentes!
- Erläutern Sie den Unterschied zwischen Prüfexperimenten und Erkundungsexperimenten in Hinblick auf ihre Funktion im Forschungsprozess!
- Erläutern Sie, welche Funktion das Milgram-Experiment im Unterschied dazu erfüllt!
- Erläutern Sie den Unterschied zwischen Pilot-Studien und Voruntersuchungen!
- Erläutern Sie, warum naturwissenschaftliche Gesetzesaussagen nicht isoliert überprüft werden können!
- Erläutern Sie, wie man zur Geltungsprüfung der Theorien stattdessen vorgehen muss!
- Erläutern Sie, was man unter einer Sachhypothese versteht! In welcher Relation steht sie zur Theorie, und wozu dient sie?
- Woran misst sich der empirische Gehalt einer Hypothese? Geben Sie ein Beispiel!
- Was versteht man unter strenger Theorienprüfung? Geben Sie Beispiele für strengere und weniger strenge Prüfung einer Theorie!
- Diskutieren Sie die wissenschaftssoziologische Problematik der Forderung nach strenger Hypothesenprüfung!
- Erläutern Sie, warum die Klärung der Terminologie eine unabdingbare Voraussetzung (strenger) Theorienprüfung ist!
- Was versteht man unter Variablenpsychologie? Geben Sie ein Beispiel!
- Worin macht sich das teilweise immer noch bestehende Theoriedefizit der Psychologie bemerkbar?
- An welche Voraussetzungen ist die experimentelle Prüfung von Theorien gebunden?
- Was bedeutet Widerspruchsfreiheit, und warum ist die Widerspruchsfreiheit eine Vorbedingung für die empirische Prüfbarkeit von Hypothesen? Geben Sie ein Beispiel!
- Zeigen Sie mittels einer Wahrheitstafel, dass man aus widerspruchshaltigen Prämissen jede beliebige Schlussfolgerung ableiten kann!
- Zeigen Sie anhand eines Beispiels, dass die Widerspruchsfreiheit einer Theorie oder Hypothese ohne Klärung der darin vorkommenden Termini nicht abschließend beurteilt werden kann!
- Erläutern Sie, was man unter der Kritisierbarkeit von Theorien ver-

steht!

- Erläutern Sie, warum die Kritisierbarkeit eine Vorbedingung für die empirische Prüfbarkeit von Hypothesen ist!
- Geben Sie Beispiele für nicht kritisierbare Theorien!
- Bestimmte Theorien sind schon aufgrund ihrer logischen Form nicht kritisierbar. Welche sind das?
- Zeigen Sie mittels der Wahrheitstafel, dass sich aus widerspruchshaltigen Theorien keine Sachhypothese ableiten lässt, welche die Theorie falsifizieren könnte!
- Zeigen Sie mittels der Wahrheitstafel, dass die Bauernregel "Kräht der Hahn auf dem Mist, ändert sich das Wetter, oder es bleibt, wie es ist" eine Tautologie ist!
- Zeigen Sie mittels der Wahrheitstafel, dass Zirkelschlüsse stets tautologisch sind!
- Erläutern Sie die Bedeutung des Terminus ‚formal kritisierbar‘!
- Zeigen Sie anhand eines Beispiels, dass die Kritisierbarkeit einer Theorie ohne Klärung der darin vorkommenden Termini nicht abschließend beurteilt werden kann!
- Erläutern Sie das Prinzip der Realitätsverdoppelung am Beispiel der naiven Triebtheorie der Aggression!
- Skizzieren Sie die Triebtheorie der Aggression nach Konrad Lorenz und erläutern Sie, inwiefern es sich dabei nicht um eine bloße Verdoppelung der Realität handelt!
- Erläutern Sie, welche Argumente Franck gegen die Schlüssigkeit der Isolationsexperimente vorbringt, mittels derer Lorenz die Staubbarkeit der Aggression belegen wollte!
- Beschreiben Sie die Untersuchung von Sipes und diskutieren Sie, inwiefern die Untersuchungsergebnisse als Falsifikation der von Lorenz behaupteten Auslebbarkeit der Aggression gelten können!
- Worin ist trotz des Scheiterns seiner Theorie die große Leistung von Konrad Lorenz zu sehen?
- Was versteht man unter der Kontingenz der Sachhypothese?
- Begründen Sie, warum die Kontingenz der Sachhypothese eine Vorbedingung für die empirische Prüfung von Theorien ist!
- Mittels welcher Testfragen kann man einen begründeten Verdacht schöpfen, dass der Zusammenhang zwischen UV und AV nicht empirisch, sondern strukturell ist? Geben Sie ein Beispiel!
- In welchen Fällen muss mit der Gefahr eines strukturellen Zusammenhanges zwischen UV und AV besonders gerechnet werden? Warum ist das so?
- Zeigen Sie anhand eines Beispiels, dass die Kontingenz einer Hypothese ohne Klärung der darin vorkommenden Termini nicht abschließend beurteilt werden kann!
- Wie kann der Nachweis erbracht werden, dass es sich bei einer gegebenen Hypothese um keine kontingente Hypothese handelt, und was ist dazu erforderlich?
- Zeigen Sie, dass es sich bei der oben diskutierten Hypothese um einen strukturellen Zusammenhang handelt: "Wenn jemand in einer winlose-Situation nicht nachgibt, wird er versuchen, sich durchzusetzen".
- Was versteht man unter der Operationalisierbarkeit einer Hypothese?

- Wann ist die Operationalisierbarkeit psychologischer Variablen eher unproblematisch, und wann erweist sie sich als schwieriger? Geben Sie Beispiele!
- Welche Voraussetzungen müssen gegeben sein, damit eine Hypothese (korrekt) operationalisiert werden kann? Geben Sie ein Beispiel oder Gegenbeispiel!
- Welche empirischen Voraussetzungen sind erforderlich, um NS, US und UR durch Glockenton, Futter und Speichelfluss operationalisieren zu können?
- Welche Zusatzannahmen sind erforderlich, um die Wahrnehmung des Stimulusunterschiedes in Experimenten zum Weber'schen Gesetz durch die Auskünfte der Vpn operationalisieren zu können?
- Wie könnte man die Wahrnehmung des Stimulusunterschiedes im Tierversuch operationalisieren? Welche Zusatzannahmen sind dazu erforderlich?
- Inwiefern geht die Frustrations-Aggressions-Hypothese bereits über das deduktiv-nomologische Erklärungsmodell hinaus?

2.6

Das induktiv-statistische Erklärungsmodell

Bei strikter Geltung des Weber'schen Gesetzes und einer relativen Unterschiedsschwelle von $c = 1/50 = 4/200$ würden wir als Ausgang unseres oben dargestellten Experimentes erwarten, dass alle Stimulusdifferenzen bis einschließlich $3 \times S/200$ von den Probanden nicht wahrgenommen werden, während alle Stimulusdifferenzen ab $4 \times S/200$ und darüber wahrgenommen werden. Stellen wir die relativen Häufigkeiten der wahrgenommenen Stimulusdifferenzen in einem Diagramm dar, so müsste sich somit das in Abb. 2.6.1 dargestellte Bild ergeben.

Abb. 2.6.1: Weber'sches Gesetz als deterministisches Naturgesetz

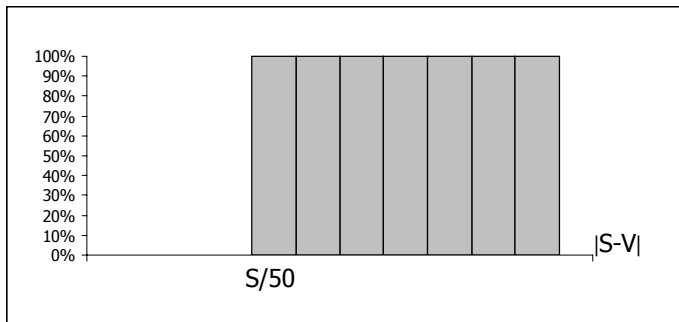
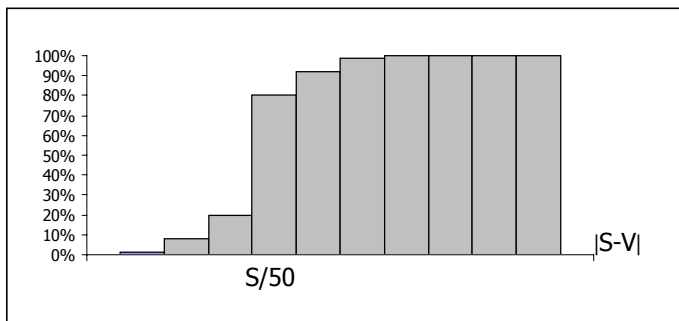


Abb. 2.6.2: Weber'sches Gesetz als statistische Gesetzesaussage



Tatsächlich wird das Ergebnis aber eher so aussehen wie in Abb. 2.6.2, in der zwar sehr kleine Stimulusdifferenzen tatsächlich nie und sehr große Stimulusdifferenzen immer wahrgenommen werden. Im Bereich dazwischen besteht jedoch eine gewisse Unschärfe.

Hierbei zeigt sich erneut, was sich schon in Kapitel 2.4 verschiedentlich angedeutet hat. Psychologische Gesetze gelten nicht mit derselben Stringenz wie z.B. die Gesetze der Newton'schen Mechanik. Zumindest sind psychologische Daten mit Mess- oder Beobachtungsfehlern behaftet, so dass die Formulierung deterministischer Gesetzmäßigkeiten auf Datenebene eher unrealistisch ist. Um diesem Tatbestand Rechnung zu tragen, müssen wir das deduktiv-nomologische Erklärungsmodell abschwächen.

Wir fragen jetzt nicht mehr nach der zwingenden Ursache, warum eine bestimmte Stimulusdifferenz von einem Probanden wahrgenommen wird oder nicht, sondern wir fragen nach den Bedingungen, unter denen wir damit rechnen müssen, dass die Stimulusdifferenz *wahrscheinlich* wahrgenommen wird.

An die Stelle des deduktiv-nomologischen Erklärungsmodells tritt damit das induktiv-statistische Erklärungsmodell (Hempel, 1965), in dem die deterministische Gesetzesaussage $\forall x. A(x) \rightarrow B(x)$ durch eine statistische Gesetzesaussage ersetzt ist, welche die Wahrscheinlichkeit (engl. *probability*) angibt, mit welcher wir mit dem Eintreffen des zu erklärenden Sachverhaltes B rechnen können, wenn die Randbedingung A gegeben ist:

$A(N)$	Randbedingung
$\text{prob}\{B(x) A(x)\} = p$	statistische Gesetzesaussage
$B(N)$	Explanandum

Zum Beispiel:

$$\frac{|S-V| > c \times S}{\text{prob}\{w(S-V) \mid |S-V| > c \times S\}} = \text{groß}$$

$$w(S-V)$$

"Es ist mit großer Wahrscheinlichkeit erwartbar, dass der Stimulusunterschied wahrgenommen wird, da er dem Betrage nach größer als $c \times S$ ist".

Infolge der Abschwächung des Erklärungsmodells ist der Schluss von den Randbedingungen und der (statistischen) Gesetzesaus-

sage auf das Explanandum jetzt aber kein logisch zwingender Schluss mehr, sondern das Eintreten von $B(N)$ ist unter den gegebenen Randbedingungen lediglich plausibel. $B(N)$ ist erwartbar, wenn $A(N)$ zutrifft und p groß ist. Man sagt auch: Das Explanandum $B(N)$ wird durch die statistische Gesetzesaussage induktiv gestützt.

An die Stelle deterministischer Hypothesen, welche die AV als naturnotwendige Wirkung der UVn darstellen, sind im induktiv-statistischen Erklärungsmodell probabilistische Hypothesen getreten, welche den kontingenten Zusammenhang zwischen UV und AV quantifizieren und – unter Voraussetzung der Zufälligkeit der AV – auf einer Absolutskala zwischen Null (unmögliches Ereignis) und Eins (notwendiges Ereignis) messen.

2.6.1 Zufall und Wahrscheinlichkeit

Die Axiome von Kolmogoroff

In der mathematischen Statistik stellt man die wesentlichen Eigenschaften des Wahrscheinlichkeitsbegriffs durch die Axiome von Kolmogoroff (1933) dar, aus denen die gesamte Wahrscheinlichkeitsrechnung herleitbar ist:

1. Die Wahrscheinlichkeit $\text{prob}\{A\}$ eines Ereignisses A bei einem Zufallsexperiment ist eine eindeutig bestimmte nichtnegative reelle Zahl, die höchstens gleich 1 sein kann:

$$0 \leq \text{prob}\{A\} \leq 1$$

2. Die Wahrscheinlichkeit eines sicheren Ereignisses S , das bei einem Zufallsexperiment stets eintritt, ist gleich 1:

$$\text{prob}\{S\} = 1$$

3. Schließen sich zwei Ereignisse A und B bei einem Zufallsexperiment gegenseitig aus, so ist die Wahrscheinlichkeit ihrer Adjunktion gleich der Summe ihrer Wahrscheinlichkeiten:

$$\text{prob}\{A \vee B\} = \text{prob}\{A\} + \text{prob}\{B\}$$

Die Bedeutung des Wahrscheinlichkeitsbegriffs

Für die Zwecke der Mathematik als Formalwissenschaft reichen diese Axiome völlig aus, nicht jedoch, um die Bedeutung des Wahrscheinlichkeitsbegriffes zu klären. Da dieselben Regeln, wie sie in den Axiomen niedergelegt sind, auch für jede relative Häufigkeit gelten, können die Axiome von Kolmogoroff noch nicht

einmal den Unterschied zwischen Wahrscheinlichkeiten und relativen Häufigkeiten erklären.

In der Psychologie müssen wir nach Einführung des induktiv-statistischen Erklärungsmodells aber nicht nur mit Wahrscheinlichkeiten rechnen, sondern die psychologischen Theorien selbst machen nunmehr nur noch Aussagen über Wahrscheinlichkeiten. Um also zu verstehen, was diese Theorien denn eigentlich behaupten, benötigen wir ein Modell des Wahrscheinlichkeitsbegriffs. Ein solches Modell wurde von Lorenzen (1974, 1985) entwickelt.

2.6.1.1 Die konstruktive Begründung des Wahrscheinlichkeitsbegriffs¹

Lorenzen (1974, 1985) leitet den Wahrscheinlichkeitsbegriff vom Begriff des Zufalls her und definiert letzteren durch die Angabe von Konstruktionsprinzipien für Zufallsgeneratoren, wobei zwischen diskreten und kontinuierlichen Zufallsgeneratoren unterschieden wird.

Zufällige Ereignisse

Dass das Eintreten eines Sachverhaltes ein zufälliges Ereignis ist, heißt dann nichts anderes, als dass er unter Benutzung eines Zufallsgenerators herbeigeführt worden ist oder so behandelt werden kann, als ob er so herbeigeführt worden wäre, "als ob Gott würfle". Die Anwendung eines Zufallsgenerators bezeichnen wir als Zufallsexperiment.

Das klassische Beispiel, an dem man den Gebrauch von Wahrscheinlichkeitsbehauptungen einzuüben pflegt, ist der Würfel, bei dem jeder der als Wurfergebnis möglichen Augenzahlen 1, 2, ... 6 dieselbe Wahrscheinlichkeit $1/6$ zugeordnet wird. Dass die Wahrscheinlichkeit, eine bestimmte Augenzahl zu würfeln, $1/6$ ist, wird dadurch begründet, dass ein regelmäßiger, aus homogenem Material bestehender Würfel eben so konstruiert ist, dass nach keinem Wissen eine Voraussage darüber gemacht werden kann, auf welche Seite der Würfel fällt. Nicht nur, dass ein solches Wissen nicht verfügbar ist, sondern man weiß positiv, dass eine solche Voraussage nicht möglich ist.

Diskrete
Zufallsgeneratoren

1. Dieser Abschnitt kann ohne Verlust der Kontinuität übersprungen werden.

Das bedeutet, dass ein solcher Würfel ein (diskreter) Zufalls-generator ist, der den folgenden Konstruktionsprinzipien genügt:

1. *Eindeutigkeit*: Jede korrekte Anwendung des Zufalls-generators (jeder Versuch) ergibt als Resultat genau eines von endlich vielen Elementarereignissen E_1, \dots, E_m .
2. *Ununterscheidbarkeit*: Mit keinem Wissen lässt sich ein Grund angeben, der eines der Elementarereignisse vor einem anderen auszeichnet.
3. *Wiederholbarkeit*: Nach jedem Versuch ist das Gerät wieder in demselben Zustand wie vorher

Wahrscheinlichkeit
als Quantifizierung
der Kontingenzt

Der Wahrscheinlichkeitsbegriff wird dann zur Quantifizierung der Kontingenzt zufälliger Ereignisse zwischen den beiden Polen der Unmöglichkeit und der Sicherheit eingeführt. Die Wahrscheinlichkeit soll also beschreiben, wie sehr ein zufälliges Ereignis eher zu den unmöglichen oder zu den sicheren Ereignissen tendiert.

Wegen des Prinzips der Wiederholbarkeit macht es Sinn, die Kontingenzt zufälliger Ereignisse durch eine Häufigkeit zu beschreiben: Unmögliche Ereignisse (U) treten bei noch so langen Versuchsreihen nie ein; sichere Ereignisse (S) treten immer ein. Je mehr also ein Ereignis zur Unmöglichkeit tendiert, desto seltener ist es; je mehr es zur Sicherheit tendiert, desto häufiger ist es (und umgekehrt).

Da die Wahrscheinlichkeit eines Ereignisses eine Eigenschaft des Zufallsexperimentes beschreiben soll, als dessen Ergebnis es eintreten kann, und nicht eine Eigenschaft der jeweiligen Versuchsreihe, welche wir aktuell durchgeführt haben, ist es zudem sinnvoll, die Wahrscheinlichkeit durch eine *relative* Häufigkeit zu beschreiben, so dass

$$\text{prob}\{U\} = 0 \leq \text{prob}\{A\} \leq \text{prob}\{S\} = 1.$$

Der Laplace'sche
Wahrscheinlichkeits-
begriff

Wegen des Prinzips der Eindeutigkeit ist bei diskreten Zufalls-generatoren die Adjunktion der Elementarereignisse $E_1 \vee \dots \vee E_m$ ein sicheres Ereignis mit

$$\text{prob}\{E_1 \vee \dots \vee E_m\} = 1,$$

wobei sich die Elementarereignisse wechselseitig ausschließen, so dass aufgrund der Rechenregeln für relative Häufigkeiten

$$\text{prob}\{E_1 \vee \dots \vee E_m\} = \text{prob}\{E_1\} + \dots + \text{prob}\{E_m\}.$$

Wegen der Ununterscheidbarkeit der Elementarereignisse muss ihnen aber allen dieselbe Wahrscheinlichkeit zukommen, so dass

$$\text{prob}\{E_1\} = \dots = \text{prob}\{E_m\} = \frac{1}{m},$$

und, wenn die A auf k der Elementarereignisse zutrifft, ergibt sich der Laplace'sche Wahrscheinlichkeitsbegriff

$$\text{prob}\{A\} = \frac{k}{m} = \frac{\text{Anzahl der "günstigen" Elementarereignisse}}{\text{Anzahl der möglichen Elementarereignisse}} \quad [13]$$

Um schließlich noch zu zeigen, dass die so definierte Wahrscheinlichkeit tatsächlich die gewünschte Quantifizierung der Kontingenz leistet, verweist Lorenzen auf das Gesetz der großen Zahlen, wonach die relative Häufigkeit des Eintretens eines zufälligen Ereignisses A mit wachsendem Stichprobenumfang gegen die oben definierte Wahrscheinlichkeit des Ereignisses konvergiert (vgl. Abschnitt 2.6.1.2). Etwas ungenau ausgedrückt: Wahrscheinlichkeit ist die relative Häufigkeit (zufälliger Ereignisse) auf Dauer.

Als Beispiel für einen kontinuierlichen Zufallsgenerator betrachten wir das Glücksrad (vgl. Abb. 2.6.3).

Kontinuierliche Zufallsgeneratoren

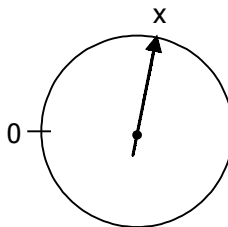


Abb. 2.6.3: Das Glücksrad. Länge der Kreislinie = 1 Einheit; x = Abstand des durch den Zeiger markierten Punktes auf der Kreislinie von einem beliebig markierten Nullpunkt

1. Anders als bei diskreten Zufallsgeneratoren gibt es hier keine Elementarereignisse, sondern einen kontinuierlichen Wertebereich, der auf unendlich viele Weisen in Intervalle eingeteilt werden kann.
2. Für jede mögliche Einteilung der Kreislinie in m Intervalle der Breite $\Delta = \frac{1}{m}$ und für jedes $m = 1, 2, 3, \dots$ sind die zum selben m gehörigen Intervalle ununterscheidbar.
3. Nach jedem Versuch ist das Gerät wieder in demselben Zustand wie vorher (Wiederholbarkeit).

Wegen der Ununterscheidbarkeit der Intervalle erhält dann jedes der m Intervalle $[x_i - \Delta/2, x_i + \Delta/2]$ die gleiche Wahrscheinlichkeit

$1/m$ zugeordnet (und zwar aus denselben Gründen, aus denen die Elementarereignisse diskreter Zufallsgeneratoren gleich wahrscheinlich sind). Da die Kreislinie zur Länge 1 normiert ist, ist die Wahrscheinlichkeit eines jeden beliebigen Intervalls gleich seiner Länge. Folglich gilt für jede Zahl x mit $0 \leq x \leq 1$ die Beziehung $\text{prob}\{X \leq x\} = x$, und die kontinuierliche Zufallsvariable X folgt einer Rechtecksverteilung (vgl. Abb. 2.6.4) mit der Verteilungsfunktion

$$\text{prob}\{X \leq x\} = F(x) = \int_{-\infty}^x f(v)dv = \begin{cases} 0 & \text{für } x < 0 \\ x & \text{für } 0 \leq x \leq 1 \\ 1 & \text{für } x > 1 \end{cases}$$

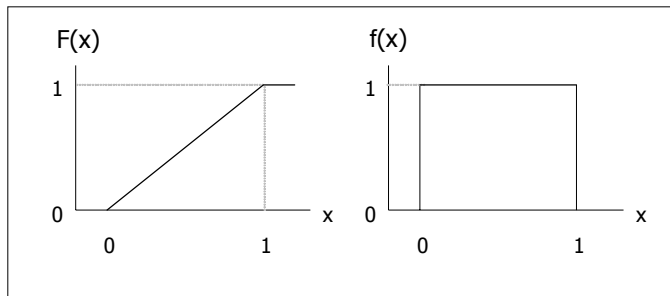
und der Wahrscheinlichkeitsdichte

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} 1 & \text{für } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Für jedes Intervall $[a, b]$ mit $a \geq 0$ und $b \leq 1$ ist die Wahrscheinlichkeit des Ereignisses $B \equiv (a \leq X \leq b)$ gleich der Länge des Intervalls

$$\text{prob}\{a \leq X \leq b\} = \int_{-\infty}^b f(v)dv - \int_{-\infty}^a f(v)dv = b - a \quad [14]$$

Abb. 2.6.4: Verteilungsfunktion und Wahrscheinlichkeitsdichte der Rechteckverteilung



Nicht-elementare
zufällige Ereignisse

Als zufällig können nicht nur jene Ereignisse gelten, die als unmittelbares Ergebnis eines Zufallsexperimentes eintreten können, wie z.B. die Elementarereignisse diskreter Zufallsgeneratoren oder die Ereignisse A und B in den Gleichungen [13] und [14].

Zufällig sind auch all jene Ereignisse, die durch Operationen der Vergrößerung ($A \vee B$), der Produktbildung ($A \wedge B$) oder der Relativierung ($A|B$) zufälliger Ereignisse entstehen, die bei der Ausfüh-

zung eines oder mehrerer Zufallsexperimente oder auch bei der wiederholten Ausführung desselben Zufallsexperimentes eintreten können sowie all jene Ereignisse, die zwar nicht durch Anwendung eines Zufallsgenerators zustande kommen, deren Entstehungsbedingungen sich jedoch auf die Anwendung eines oder mehrerer Zufallsgeneratoren abbilden lassen.

So verletzt z.B. das Ziehen aus einer Urne ohne Zurücklegen das Prinzip der Wiederholbarkeit und ist deshalb selbst kein Zufallsgenerator. Andererseits lässt es sich jedoch auf das Ziehen mit Zurücklegen aus mehreren, unterschiedlich gefüllten Urnen abbilden, wobei das Ergebnis der vorangegangenen Züge darüber entscheidet, aus welcher Urne der jeweils nächste Zug zu tätigen ist (vgl. Abb. 2.6.5).

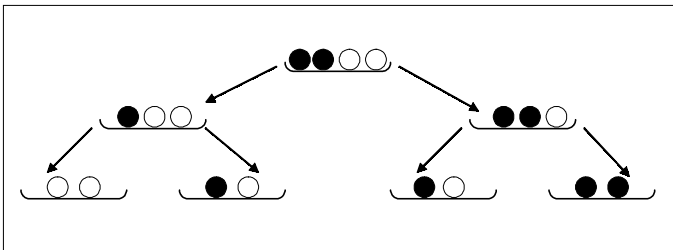


Abb. 2.6.5: Ziehen ohne Zurücklegen

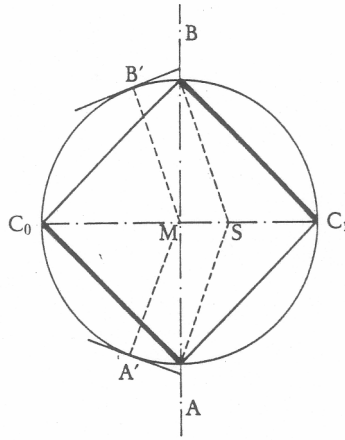
Auch unregelmäßige Würfel sind selbst keine Zufallsgeneratoren, da sie das Prinzip der Ununterscheidbarkeit verletzen. Sie können jedoch auf die Vergrößerung eines kontinuierlichen Zufallsgenerators abgebildet werden, so dass man das Werfen eines "falschen" Würfels ebenfalls so behandeln kann, "als ob" es durch Anwendung eines Zufallsgenerators zustande gekommen wäre.

Zwecks besserer Anschaulichkeit reduziert Lorenzen (1974) das Problem auf zwei Dimensionen: Statt eines Würfels wird eine quadratische Säule geworfen, deren Schwerpunkt S vom Mittelpunkt M verschieden ist (vgl. Abb. 2.6.6).

Unregelmäßige
Würfel

Wir denken uns die Säule $C_0 A C_1 B$ in einem Zylinder eingeschlossen. Der Umfang des Zylinders sei zu 1 normiert. Die Säule werde so geworfen, dass der (gedachte) Zylinder mit gleicher Wahrscheinlichkeit innerhalb gleich langer Stücke seines Umfanges senkrecht ohne Drehimpuls auf den Tisch fällt. Ohne den Zylinder fällt die quadratische Säule dadurch auf eine der vier Seiten. Wir fragen nach der Wahrscheinlichkeit, dass C_0 oben liegen wird.

Abb. 2.6.6: Die quadratische Säule (nach Lorenzen, 1974, 215)



Werden A' bzw. B' auf dem Kreisumfang dadurch bestimmt, dass $A'M$ bzw. $B'M$ parallel zu AS bzw. BS sind, dann ist die Länge des Kreisbogens $A'C_1B'$ die gesuchte Wahrscheinlichkeit.

Ist die Lage des Schwerpunktes bekannt, so lässt sich diese Wahrscheinlichkeit $\text{prob}\{C_0\} = p > 0$ berechnen. Ohne vorherige Bestimmung des Schwerpunktes findet man durch die relative Häufigkeit h_n in genügend langen Versuchsreihen der Länge n nach dem Gesetz der großen Zahlen die gut gestützte Hypothese $p = h_n$.

2.6.1.2 Exkurs: Das Gesetz der großen Zahlen¹

Das Gesetz der großen Zahlen

Es sei A ein Ereignis, das bei einem Zufallsexperiment die Wahrscheinlichkeit p besitzt ($0 < p < 1$). Das Experiment werde n mal ausgeführt. Dann entspricht jeder Ausführung des Experimentes ($i = 1, \dots, n$) eine unabhängige Realisation (x_i) einer Bernoulli-Variablen mit

$$x = \begin{cases} 1 & \text{falls das Ereignis } A \text{ eintritt} \\ 0 & \text{sonst} \end{cases}$$

und der Wahrscheinlichkeitsfunktion

1. Dieser Abschnitt kann ohne Verlust der Kontinuität übersprungen werden.

$$f(x) = \begin{cases} p & \text{für } x = 1 \\ 1 - p & \text{sonst} \end{cases}$$

$X_o = \sum_{i=1}^n X_i$ bezeichne die (absolute) Häufigkeit des Eintreffens von A.

$h_n = \frac{X_o}{n}$ bezeichne die relative Häufigkeit. Weiterhin sei eine beliebig kleine positive Zahl ϵ vorgegeben. Dann gilt

$$\lim_{n \rightarrow \infty} \text{prob}\{|h_n - p| < \epsilon\} = 1.$$

Mit wachsendem n strebt die Wahrscheinlichkeit, dass die relative Häufigkeit des Eintreffens von A um mehr als eine beliebig kleine positive Zahl ϵ von p abweicht, gegen Null.

Zum Beweis des Gesetzes benötigen wir noch eine Reihe von Voraussetzungen:

Die relative Häufigkeit h_n ist eine erwartungstreue Schätzfunktion für den Parameter p der Bernoulli-Verteilung, so dass

Die Erwartungstreue der relativen Häufigkeit

$$E(h_n) = p$$

Beweis: Wegen $h_n = \frac{X_o}{n} = \frac{1}{n} \sum_{i=1}^n X_i$ ist der Erwartungswert

$$E(h_n) = \frac{1}{n} E\left(\sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i)$$

Die X_i haben aber alle dieselbe Bernoulli-Verteilung mit

$$E(X_i) = \sum_{x=0}^1 xf(x) = p$$

so dass

$$E(h_n) = \frac{1}{n} np = p$$

Die Varianz der relativen Häufigkeit beträgt

Die Wirksamkeit der relativen Häufigkeit

$$\sigma^2(h_n) = \frac{p(1-p)}{n}$$

Beweis: Wegen $h_n = \frac{1}{n} \sum_{i=1}^n X_i$ ist

$$\sigma^2(h_n) = \frac{1}{n^2} \sigma^2\left(\sum_{i=1}^k X_i\right)$$

und aufgrund der Unabhängigkeit der X_i ist

$$\sigma^2\left(\sum_{i=1}^k X_i\right) = \sum_{i=1}^k \sigma^2(X_i)$$

so dass

$$\sigma^2(h_n) = \frac{1}{n^2} \sum_{i=1}^k \sigma^2(X_i)$$

Die X_i haben aber alle dieselbe Bernoulli-Verteilung mit $E(X_i) = p$ und

$$E(X_i^2) = \sum_{x=0}^1 x^2 f(x) = p$$

Daher ist $\sigma^2(X_i) = E(X_i^2) - E(X_i)^2 = p - p^2 = p(1 - p)$. Also ist die Summe

$$\sum_{i=1}^n \sigma^2(X_i) = np(1-p)$$

und für die Varianz der relativen Häufigkeit folgt

$$\sigma^2(h_n) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}$$

Die Ungleichung von
Tschebyscheff

Sei X eine beliebige Zufallsvariable mit endlichem Erwartungswert $E(X)$ und endlicher Varianz σ^2 , so ist für jede positive Zahl c die Ungleichung

$$\text{prob}\{|X - E(X)| \geq c\sigma\} \leq \frac{1}{c^2}$$

erfüllt.

*Beweis:*¹ Unter den getroffenen Voraussetzungen ist $Y = (X - E(X))^2$ eine zufällige Variable, die nur nichtnegative Werte annehmen kann und die den endlichen Erwartungswert $E(Y) = E\{(X - E(X))^2\} = \sigma^2$ besitzt. Die Wahrscheinlichkeitsdichte von Y sei mit $g(y)$ bezeichnet. Folglich gilt für jede positive Zahl a

1. Wir beweisen die Ungleichung für stetige Zufallsvariablen. Für diskrete Zufallsvariablen erfolgt der Beweis analog und sei dem Leser als Übungsaufgabe überlassen.

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} yg(y)dy = \int_{-\infty}^a yg(y)dy + \int_a^{\infty} yg(y)dy \\ &\geq \int_a^{\infty} yg(y)dy \geq a \int_a^{\infty} g(y)dy = a \text{ prob}\{Y \geq a\} \end{aligned}$$

so dass wir zunächst die Ungleichung

$$\text{prob}\{Y \geq a\} \leq \frac{E(Y)}{a}$$

erhalten. Einsetzen für $Y = (X - E(X))^2$ und $a = c^2\sigma^2$ ergibt:

$$\text{prob}\{(X - E(X))^2 \geq c^2\sigma^2\} \leq \frac{\sigma^2}{c^2\sigma^2} = \frac{1}{c^2}$$

Diese Ungleichung ist der Tschebyscheffschen Ungleichung äquivalent.

Beweis des Gesetzes der großen Zahlen: Einsetzen von $X = h_n$ und entsprechend

Der Beweis des Gesetzes der großen Zahlen

$$E(X) = p$$

und

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

in die Ungleichung von Tschebyscheff ergibt

$$\text{prob}\left\{|h_n - p| \geq c \sqrt{\frac{p(1-p)}{n}}\right\} \leq \frac{1}{c^2}$$

O.B.d.A. setzen wir

$$\varepsilon = c \sqrt{\frac{p(1-p)}{n}}$$

so dass

$$\frac{1}{c} = \frac{\sqrt{\frac{p(1-p)}{n}}}{\varepsilon}$$

und

$$\frac{1}{c^2} = \frac{p(1-p)}{n\varepsilon^2}$$

Also ist

$$\text{prob}\{|h_n - p| \geq \varepsilon\} \leq \frac{p(1-p)}{n\varepsilon^2}$$

und der Grenzwert

$$\lim_{n \rightarrow \infty} \text{prob}\{|h_n - p| \geq \varepsilon\} = 0.$$

Diese Gleichung ist dem Gesetz der großen Zahlen äquivalent.

2.6.2 Statistische Modellbildung

Wahrscheinlichkeiten sind idealisierte relative Häufigkeiten, und zwar in zweierlei Hinsicht:

1. indem die Zufälligkeit des Ereignisses unterstellt wird, und
2. indem man die Anzahl der Beobachtungen, auf welchen diese beruht, in einer Art Gedankenexperiment gegen unendlich gehen lässt.

Die Unterstellung der Zufälligkeit wirft die Frage auf, an welcher Stelle des Forschungsprozesses der Zufall ins Spiel kommt. Dabei lassen sich verschiedene Möglichkeiten unterscheiden, die wir wieder anhand des Weber'schen Gesetzes erläutern können.

Die erste Art der Modellbildung unterstellt, dass die Wahrnehmung der Stimulusdifferenz selbst ein Zufallsprozess ist (wobei die Wahrscheinlichkeit außerhalb des Schwellenbereiches allerdings gegen Null bzw. Eins tendiert).

Das Weber'sche Gesetz besagt dann, dass die Wahrscheinlichkeit, mit welcher eine Stimulusdifferenz wahrgenommen wird, eine Funktion der Stimulusdifferenz selbst ist, bzw. genauer: eine Funktion der relativen Unterschiedschwelle (c), des Standardreizes (S) und des Vergleichsreizes (V) (vgl. Abb. 2.6.7):

$$\text{Prob}\{w(S-V)\} = f(c, S, V).$$

Diese Art von Modellen bezeichnet man als stochastische Modelle. Beispiele für stochastische Modellbildungen sind die Latent-Trait- und Latent-Class-Modelle in der Psychometrie und/oder das lineare logistische Modell in der Statistik.

Stochastische Modelle

Fehlermodelle mit fester Effektgröße

Die zweite Art der Modellbildung hält an der deterministischen Geltung des Weber'schen Gesetzes fest, postuliert aber, dass die Be-

obachtungen durch zufällige Störeinflüsse und/oder Beobachtungsfehler verzerrt werden, so dass die Messwerte der (absoluten) Unterschiedsschwelle zufällig variieren, d.h. eine Zufallsvariable (Δ) darstellen.

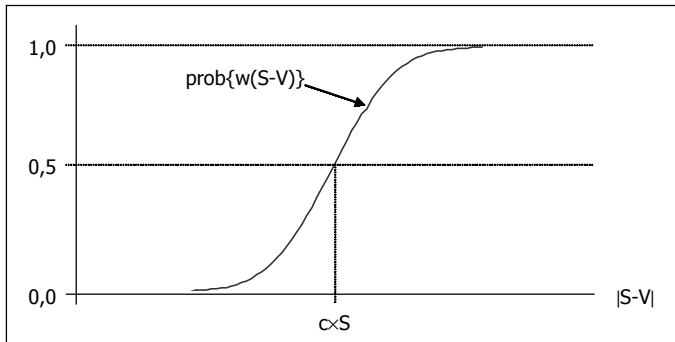


Abb. 2.6.7: Formulierung des Weber'schen Gesetzes als stochastisches Modell

Zum Weber'schen Gesetz tritt also ein Zufallsfehler (F) hinzu, der sich im Durchschnitt über alle möglichen Beobachtungen ausgleicht ($E(F) = 0$), so dass das Gesetz in der Form

$$\Delta = c \cdot S + F$$

geschrieben werden kann und nur noch im Durchschnitt über alle möglichen Beobachtungen exakt gilt (vgl. Abb. 2.6.8).

$$E(\Delta) = c \cdot S.$$

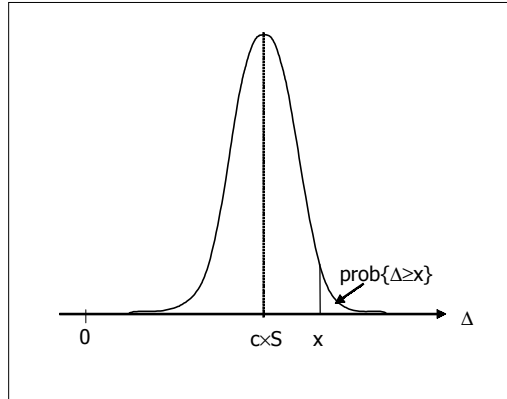
Die Gesetzesaussage wird nur für das mittlere Δ behauptet: Im statistischen Durchschnitt ist Δ gleich $c \cdot S$.

Diese Art von Modellen bezeichnet man als Fehlermodelle mit fester Effektgröße. Beispiele sind die klassische Testtheorie in der Psychometrie und die Varianzanalyse mit fester Effektgröße (fixed effect) in der Statistik. Feste Effektgröße bedeutet, dass der Effekt der UV über die V_{pn} hinweg konstant ist.

Während wir es bei den bisher dargestellten Modellen mit *Zufallsexperimenten* zu tun hatten, gehen wir bei der dritten Art der Modellbildung von einer *Zufallsauswahl* aus.

- Ein Zufallsexperiment liegt vor, wenn die beobachteten Ereignisse selbst das Produkt eines Zufallsprozesses sind (z.B. Würfel).

Abb.2.6.8: Formulierung des Weber'schen Gesetzes als Fehlermodell mit fester Effektgröße



- Eine Zufallsauswahl liegt vor, wenn aus der Menge der in Frage kommenden Gegenstände mittels Zufallsgenerator einige ausgewählt und tatsächlich beobachtet werden (z.B. Urne).

Fehlermodelle mit zufälliger Effektgröße

Mit einer solchen Zufallsauswahl haben wir es z.B. dann zu tun, wenn wir annehmen, dass das Weber'sche Gesetz

- (im Unterschied zu den stochastischen Modellen) für jeden einzelnen Probanden exakt gilt, und
- dass unsere Beobachtungen (im Unterschied zu den Fehlermodellen) auch nicht durch einen Zufallsfehler überlagert werden, sondern
- dass die relative Unterschiedsschwelle (c_v) jedoch für jeden Probanden (v) geringfügig anders ist, und
- dass die Probanden zufällig ausgewählt wurden.

In diesem Falle ist die relative Unterschiedsschwelle selbst eine (über die Probanden hinweg) zufällig verteilte Variable (C), und das Weber'sche Gesetz ist in der Form

$$\Delta_v = c_v \times S$$

bzw.

$$D = C \times S$$

zu präzisieren. Auch die absolute Unterschiedsschwelle ist dann nämlich eine zufällige Variable (D), und im Durchschnitt über die Probanden hinweg gilt:

$$E(D) = E(C) \times S.$$

Definieren wir $\Delta S = E(D)$ und $\mu_c = E(C)$, so lässt sich das Weber'sche Gesetz nun in der Form

$$\Delta S = \mu_c \times S$$

darstellen, wird in dieser (scheinbar) deterministischen Form aber nur noch für die über die Probanden hinweg gemittelten Unterschiedsschwellen behauptet.

Mit der Definition $F = C \times S - \Delta S$ lässt sich das Weber'sche Gesetz schließlich wieder in der Form eines Fehlermodells darstellen (vgl. Abb. 2.6.9)

$$D = \mu_c \times S + F.$$

Anders als oben beschreibt die Zufallsvariable F aber hier keine Störeinflüsse oder Beobachtungsfehler, sondern die Abweichung der individuell verschiedenen Unterschiedsschwellen (Δ_v) von der mittleren Unterschiedsschwelle ΔS . Man spricht daher von einem Fehlermodell mit zufälliger Effektgröße.

Ein Beispiel für diese Art der Modellbildung ist die Varianzanalyse mit zufälliger Effektgröße (random effect) in der Statistik. Zufällige Effektgröße bedeutet, dass der Effekt der UV eine über die Vpn hinweg variierende zufällige Variable ist.

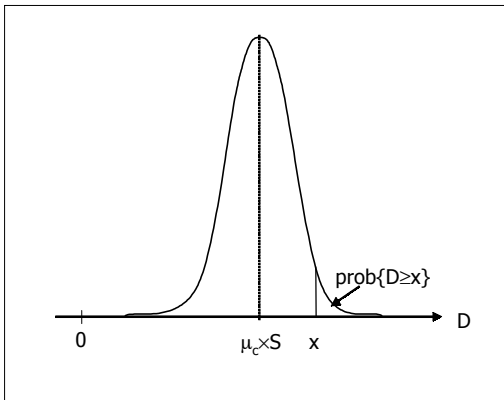


Abb. 2.6.9: Formulierung des Weber'schen Gesetzes als Fehlermodell mit zufälliger Effektgröße

Die kontrafaktische Unterstellung der Wiederholbarkeit

Die Problematik der Wahrscheinlichkeitsaussagen in der Psychometrie¹

Zufalls- und Wahrscheinlichkeitsbegriff sind wesentlich an die Wiederholbarkeit des Zufallsexperimentes gebunden (vgl. Kapitel 2.6.1.1). Sowohl in der klassischen Testtheorie als auch in der Item-Response-Theorie kann diese Voraussetzung jedoch nur *kontrafaktisch* unterstellt werden, indem wir so tun, *als ob* dieselbe V_p denselben Test bzw. dasselbe Item beliebig oft bearbeiten könnte, ohne dass die Erfolgchancen der V_p sich dadurch in irgendeiner Weise veränderten.

Klassische Testtheorie

Für die klassische Testtheorie (vgl. Kapitel 2.1.3 - 2.1.4) hat dies insofern keine weiteren Konsequenzen, als an Stelle der individuellen Scorevariablen X_{vt} , deren Zufälligkeit nur kontrafaktisch unterstellt werden kann, ohnedies bloß die Verteilung der Score- und True-Score-Variablen X_{ot} und T_{ot} in der Personenpopulation untersucht wird.

Latent-Class-Analyse

In der Latent-Class-Analyse (vgl. Kapitel 2.4.8) wird die Verletzung der Zufälligkeit der individuellen Antwortvariablen X_{vi} durch die Äquivalenz aller derselben Klasse (g) angehörigen Probanden geheilt, so dass $\forall_{v \in g} X_{vi} \equiv X_{gi}$ und die Wiederholbarkeit über die V_p n hinweg gewährleistet ist.

Latent-Trait-Modelle

Bei den Latent-Trait-Modellen (vgl. Kapitel 2.4.6) resultiert die mangelnde Zufälligkeit der individuellen Antwortvariablen X_{vi} dagegen darin, dass auch die Antwortmatrix \mathbf{X} einer gegebenen Stichprobe von n V_p n und k Items nur kontrafaktisch als zufälliges Ereignis behandelt werden kann. Die (unbedingte) Likelihood der Antwortmatrix (vgl. Gleichung [8]) stellt daher keine Wahrscheinlichkeit dar, was schließlich zum Versagen der Maximum-Likelihood-Methode führt.

Die in Gleichung [9] dargestellte bedingte Likelihood beschreibt dagegen sehr wohl die Wahrscheinlichkeit eines wiederholbaren Ereignisses, und zwar des Ereignisses, bei gegebenen Randvektoren (x_{10}, \dots, x_{n0}) eine bestimmte Antwortmatrix \mathbf{X}

1. Dieser Abschnitt kann ohne Verlust der Kontinuität übersprungen werden.

zu beobachten. Ihre Existenz ist allerdings an die Voraussetzung logistischer Itemcharakteristikfunktionen (Rasch-Modell) gebunden.

Auch die in Gleichung [10] dargestellte marginale Likelihood beschreibt die Wahrscheinlichkeit eines wiederholbaren Ereignisses, nämlich in einer Zufallsstichprobe von n Vpn die Antwortmatrix \mathbf{X} zu beobachten.

Gegenüber der CML-Methode hat die MML-Methode den Vorteil, dass sie nicht an eine bestimmte Form der Itemcharakteristiken gebunden ist. Dieser Vorteil wird jedoch durch den Nachteil erkauft, dass Annahmen über die Wahrscheinlichkeitsdichte $g(\theta)$ der latenten Dimension getroffen werden müssen. Stimmen diese Annahmen nicht mit der Wirklichkeit überein, so führt dies zu einer Verzerrung der Parameterschätzung. Lediglich im Rasch-Modell ist

$$\begin{aligned} \text{ML}(\mathbf{X}) &= \prod_{v=1}^n \int_{-\infty}^{\infty} \text{prob}\{(x_{v1}, \dots, x_{vk}) | \theta\} g(\theta) d\theta \\ &= \prod_{v=1}^n \text{prob}\{(x_{v1}, \dots, x_{vk}) | x_{v0}\} \text{prob}(x_{v0}) \end{aligned}$$

so dass keine Annahmen über $g(\theta)$ erforderlich sind und die MML-Methode zu denselben Schätzgleichungen führt wie die CML-Methode.

2.6.3 Statistische Inferenz

Im Unterschied zu relativen Häufigkeiten sind Wahrscheinlichkeiten nicht beobachtungssprachlich, sondern theoriesprachlich definiert. Daraus ergibt sich die Frage, wie man von Beobachtungen Rückschlüsse auf Wahrscheinlichkeiten ziehen kann bzw. ob und in welcher Form Aussagen über Wahrscheinlichkeiten an der Erfahrung scheitern können.

Unterschiede zwischen Häufigkeits- und Wahrscheinlichkeitsverteilungen

Empirisch beobachtbar ist nur, wie eine Variable (X) in einer endlichen Stichprobe von Beobachtungen verteilt ist. Das ist die Häufigkeitsverteilung der Variable, wie sie in der deskriptiven Statistik beschrieben wird. Die gebräuchlichsten Beschreibungsmittel sind die (relative) Häufigkeitsfunktion ($h(x_i)$), der Mittelwert (m_x), die Stichprobenvarianz (s_x^2) und die Standardabweichung einer Variable (s_x) sowie die Produktmomentkorrelation zweier Variablen (r_{xy}). Diese Kennzahlen zur Beschreibung einer Häufigkeitsvertei-

lung bezeichnet man allgemein als Statistiken und unterscheidet sie von den entsprechenden Parametern der zugrundeliegenden Wahrscheinlichkeitsverteilung, aus welcher die Stichprobe gezogen wurde: Wahrscheinlichkeitsdichte ($f(x)$), Erwartungswert (μ_x), Varianz (σ_x^2), Standardabweichung (σ_x) und Produktmomentkorrelation (ρ_{xy}).

Tab. 2.6.1: Beschreibung statistischer Häufigkeits- und Wahrscheinlichkeitsverteilungen

Statistiken zur Beschreibung empirischer Häufigkeitsverteilungen	Parameter zur Beschreibung von Wahrscheinlichkeitsverteilungen
<p>Häufigkeitsfunktion</p> $h(x_i) = h(X=x_i)$ <p>Bei diskreten Variablen stellen die x_1, x_2, \dots die diskreten Zahlenwerte dar, welche die Variable X annehmen kann</p> <p>Bei stetigen Variablen sind die x_i die Mitten von Intervallen ($x_i - \Delta/2, x_i + \Delta/2$) der Intervallbreite Δ</p>	<p>Wahrscheinlichkeitsdichte</p> $f(x)$ <p>Bei diskreten Variablen ist $f(x)$ die Wahrscheinlichkeitsfunktion mit $f(x) = \text{prob}\{X=x_i\}$ für $x=x_i$.</p> <p>Bei stetigen Variablen ist $\text{prob}\{X=x\} = 0$ und $f(x)$ ist die Dichtefunktion (s.u.)</p>
<p>Summenhäufigkeitsfunktion</p> $H(x_i) = h(X \leq x_i) = \sum_{j \leq i} h(x_j)$	<p>Verteilungsfunktion</p> $F(x) = \text{prob}\{X \leq x\}$ <p>Bei diskreten Variablen ist</p> $F(x) = \sum_{x_i \leq x} f(x_i)$ <p>Bei stetigen Variablen ist</p> $F(x) = \int_{-\infty}^x f(v) dv$ <p>Folgerungen für stetige Variablen:</p> <p>a)</p> $f(x) = \frac{dF(x)}{dx}$ <p>b) $\text{prob}\{x_i - \Delta/2 \leq X \leq x_i + \Delta/2\} =$</p> $\int_{-\infty}^{x_i + \Delta/2} f(v) dv - \int_{-\infty}^{x_i - \Delta/2} f(v) dv$ <p>c) $\text{prob}\{X=x\} =$</p> $\text{prob}\{x_i - \Delta/2 \leq X \leq x_i + \Delta/2 \Delta=0\} =$ $\int_{-\infty}^x f(v) dv - \int_{-\infty}^x f(v) dv = 0$

Mittelwert $m(X) = m_x$	Erwartungswert $E(X) = \mu_x$
Stichprobenvarianz $s^2(X) = s_x^2$ $= m\{(X - m_x)^2\} = m(X^2) - m_x^2$	Varianz $\sigma^2(X) = \sigma_x^2$ $= E\{(X - \mu_x)^2\} = E(X^2) - \mu_x^2$
Standardabweichung $s(X) = s_x = \sqrt{s_x^2}$	Standardabweichung $\sigma(X) = \sigma_x = \sqrt{\sigma_x^2}$
Kovarianz $s(X, Y) = s_{xy}$ $= m\{(X - m_x)(Y - m_y)\}$ $= m(XY) - m_x m_y$	Kovarianz $\sigma(X, Y) = \sigma_{xy}$ $= E\{(X - \mu_x)(Y - \mu_y)\}$ $= E(XY) - \mu_x \mu_y$
Produktmomentkorrelation $r(X, Y) = r_{xy} = \frac{s_{xy}}{s_x s_y}$	Produktmomentkorrelation $\rho(X, Y) = \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$
Statistiken sind Stichprobenfunktionen, z.B. $m\{g(X)\} = \frac{1}{n} \sum_{v=1}^n g(x_v)$	Parameter sind unbekannte Konstante

Zwischen diesen impliziten Parametern einer Wahrscheinlichkeitsverteilung und den entsprechenden Statistiken der Häufigkeitsverteilung besteht eine vollständige Analogie, die in Tab. 2.6.1 dargestellt ist. Ein Unterschied besteht allerdings darin, dass sich die Statistiken auf die Häufigkeitsverteilung beziehen und auch dann berechnet werden können, wenn die Variable nicht zufällig verteilt ist, während sich die Parameter auf eine Wahrscheinlichkeitsverteilung beziehen und somit die Zufälligkeit der Variable voraussetzen.

Analogie zwischen Statistiken und (impliziten) Parametern

2.6.3.1 Die Schätzung unbekannter Parameter

Da sie die Verteilung der Variable in der (unendlich großen) Grundgesamtheit aller möglichen Beobachtungen beschreiben, können die Parameter aus einer endlichen Anzahl von Beobachtungen nicht einfach berechnet werden. Sie bleiben stets theoretische Größen, die aus den Stichprobendaten lediglich geschätzt werden können. Dabei ist es zwar naheliegend, als Schätzung ei-

Schätzfunktionen

nes Populationsparameters einfach die entsprechende Stichprobenstatistik einzusetzen. Dies liefert jedoch nicht in allen Fällen eine optimale Schätzung, und bei manchen Parametern (z.B. Varianz und Standardabweichung) muss man sich daher einer korrigierten Schätzformel bedienen (vgl. Tab. 2.6.2).

Tab. 2.6.2: Einige Schätzfunktionen für unbekannte Parameter

Schätzfunktionen für unbekannte Parameter	
Erwartungswert $\hat{\mu}_x = m_x$	Schätzfunktionen sind ebenfalls Stichprobenfunktionen: $\hat{\vartheta} = u(x_1, \dots, x_n)$ Sie dienen aber nicht zur Beschreibung der jeweiligen Stichprobe, sondern zur Schätzung der unbekanntenen (theoretischen) Parameter der Wahrscheinlichkeitsverteilung, welcher die Stichprobe entstammt.
Varianz $\hat{\sigma}_x^2 = \frac{n}{n-1} s_x^2$	
Standardabweichung $\hat{\sigma}_x = s_x \sqrt{\frac{n}{n-1}}$	
Kovarianz $\hat{\sigma}_{xy} = \frac{n}{n-1} s_{xy}$	
Produktmomentkorrelation $\hat{\rho}_{xy} = r_{xy}$	

Implizite Parameter

Grundsätzlich kann zwischen impliziten und expliziten Parametern einer Wahrscheinlichkeitsverteilung unterschieden werden. Implizite Parameter sind solche, die sich durch die Momente einer Verteilung ausdrücken lassen (vgl. Tab. 2.6.1). Beispiele für implizite Parameter sind Erwartungswert, Varianz und Kovarianz sowie die daraus berechnete Standardabweichung und Produktmomentkorrelation.

Unter dem m-ten Moment einer Zufallsvariablen versteht man den Erwartungswert ihrer m-ten Potenz: $E(X^m)$. Als Produktmoment bezeichnet man entsprechend den Erwartungswert des Produktes zweier Zufallsvariablen: $E(XY)$.

Prinzipiell sind die impliziten Parameter zur Beschreibung jeder beliebigen Wahrscheinlichkeitsverteilung anwendbar; und zwar auch dann, wenn Wahrscheinlichkeitsdichte und Verteilungsfunktion unbekannt sind. Es gibt allerdings Spezialfälle von Verteilungen, bei denen bestimmte Momente nicht existieren (so besitzt z.B. die Cauchy-Verteilung weder Mittelwert noch Varianz).

Beispiele für explizite Parameter sind die Konstante c in der stochastischen Formulierung des Weber'schen Gesetzes (vgl. Kapitel 2.6.2)

Explizite Parameter

$$\text{prob}\{w(S-V)\} = f(c, S, V),$$

sowie die Item- und Personenparameter δ_i und θ_v der logistischen Itemcharakteristikfunktion (vgl. Kapitel 2.4.5)

$$\text{prob}\{X_{vi} = 1\} = \frac{e^{(\theta_v - \delta_i)}}{1 + e^{(\theta_v - \delta_i)}}$$

Allgemein bezeichnet man all jene Parameter als explizit, welche in der Funktionsgleichung der Wahrscheinlichkeitsdichte einer Zufallsvariable explizit vorkommen.

Oft können implizite Parameter durch explizite ausgedrückt werden und umgekehrt. So gilt z.B. bei der Binominalverteilung $E(X) = np$; in der Wahrscheinlichkeitsdichte der Poissonverteilung kommt der Erwartungswert μ explizit vor, und in der Dichtefunktion der Normalverteilung treten Erwartungswert μ und Standardabweichung σ explizit auf.

Gebräuchliche Verfahren zur Konstruktion von Schätzfunktionen sind die Momentenmethode und die Maximum-Likelihood-Methode.

Momentenmethode

Das älteste Verfahren ist die Momentenmethode zur Schätzung impliziter Parameter.

Dabei geht man so vor, dass man den zu schätzenden Parameter durch die Momente der Verteilung ausdrückt und die Momente $E(X)$, $E(X^2)$, $E(X^3)$, ... durch die entsprechenden Stichprobenmomente $m(X)$, $m(X^2)$, $m(X^3)$, ... ersetzt.

Wie R. A. Fisher (1921) gezeigt hat, können die mittels Momentenmethode erhaltenen Schätzfunktionen jedoch ggf. wenig wirksam und sogar auch asymptotisch wenig wirksam sein.¹

Das wichtigste Verfahren zur Gewinnung von Schätzfunktionen für die Parameter einer Wahrscheinlichkeitsverteilung ist die Maximum-Likelihood-Methode (ML-Methode), die in Sonderfällen bereits von Gauß (1880) angewendet wurde. In allgemeiner Form wurde sie von Fisher (1925) entwickelt.

Maximum-Likelihood-Methode

1. Zum Begriff der Wirksamkeit von Schätzfunktionen s.u.

Die ML-Methode dient zur Schätzung expliziter Parameter (ϑ) und berechnet die Parameterschätzungen ($\hat{\vartheta}$) so, dass die Likelihood der Stichprobenwerte

$$L\{x_1, \dots, x_n\} = \prod_{v=1}^n f(x_v) \quad [15]$$

maximal groß wird.

Die Likelihood-Funktion

Bei diskreten Zufallsvariablen ist die Likelihood nichts anderes als die Wahrscheinlichkeit der gegebenen Stichprobenwerte

$$\begin{aligned} \text{prob}\{x_1, \dots, x_n\} &= \text{prob}\{X = x_v; v = 1, \dots, n\} \\ &= \prod_{v=1}^n f(x_v) = L\{x_1, \dots, x_n\} \end{aligned}$$

Bei stetigen Zufallsvariablen ist die Wahrscheinlichkeit der Stichprobenwerte wegen $\text{prob}\{X=x\} = 0$ (vgl. Tab. 2.6.1) gleich Null. In diesem Fall gibt die Likelihood daher nicht die Wahrscheinlichkeit der Stichprobenwerte wieder. Sie ist aber proportional zu der Wahrscheinlichkeit, *in etwa* die gegebenen Stichprobenwerte zu beobachten, d.h. Stichprobenwerte zu beobachten, die in einem sehr kleinen Intervall ($\Delta \rightarrow 0$) um die gegebenen Werte liegen:

$$\begin{aligned} \text{prob}\{x_v - \frac{\Delta}{2} \leq X \leq x_v + \frac{\Delta}{2}; v = 1, \dots, n\} &= \prod_{v=1}^n \Delta f(x_v) \\ &= \Delta^n \prod_{v=1}^n f(x_v) = \Delta^n L\{x_1, \dots, x_n\} \end{aligned}$$

Varianten der ML-Methode sind die in Kapitel 2.4.6 dargestellte CML-, MML- und WML-Methode.

Die Maximierung der Likelihood-Funktion

Zur Bestimmung des Maximums der Likelihood-Funktion geht man aus rechnerischen Gründen so vor, dass man die erste Ableitung der logarithmierten Likelihood¹ gleich Null setzt und die Gleichung

$$\frac{d \ln\{L\}}{d\vartheta} = 0 \quad [16]$$

nach dem unbekanntem Parameter ϑ auflöst.

Erwartungstreue

Als Funktionen zufälliger Variablen X_1, \dots, X_n sind auch die Schätzfunktionen selbst wieder Zufallsvariablen

$$U = u(X_1, \dots, X_n),$$

1. Da der Logarithmus eine streng monoton wachsende Funktion ist, hat $\ln\{L\}$ sein Maximum an derselben Stelle wie die Likelihood selbst.

die mehr oder weniger breit um den zu schätzenden Parameter streuen. Von einer guten Schätzfunktion wird man aber erwarten, dass sie zumindest im Durchschnitt über viele Stichproben hinweg den zu schätzenden Parameter ϑ korrekt trifft, so dass

$$E(U) = \vartheta .$$

Solche Schätzfunktionen bezeichnet man als erwartungstreu.

Außerdem wird man verlangen, dass die Schätzfunktion möglichst wenig um den zu schätzenden Parameter ϑ streut, so dass **Wirksamkeit**

$$E\{(U - \vartheta)^2\} = \min .$$

Je kleiner die Fehlerstreuung $E\{(U - \vartheta)^2\}$ ist, desto wirksamer ist die Schätzfunktion.

Die Fehlerstreuung besitzt eine Untergrenze, die nicht unterschritten werden kann. Sie kommt in der nach Rao (1945) und Cramér (1946) benannten Ungleichung **Die Ungleichung von Rao-Cramér**

$$E\{(U - \vartheta)^2\} \geq \frac{1}{I(\vartheta)}$$

zum Ausdruck, worin $I(\vartheta)$ die Informationsfunktion

$$I(\vartheta) = E\left\{ \sum_{v=1}^n \left(\frac{d \ln\{f(x_v)\}}{d \vartheta} \right)^2 \right\} = n E\left\{ \left(\frac{d \ln\{f(x)\}}{d \vartheta} \right)^2 \right\}$$

bezeichnet. Schätzfunktionen, deren Fehlerstreuung diese Untergrenze nicht übersteigt, bezeichnet man als wirksamste Schätzfunktionen.

Als wirksamste Schätzfunktionen kommen nur erwartungstreue Schätzfunktionen in Betracht, und wir können zusammenfassen: **Wirksamste Schätzfunktionen**
Eine Schätzfunktion U für einen unbekanntem Parameter ϑ ist genau dann eine wirksamste Schätzfunktion, wenn sie erwartungstreu ist und ihre Varianz gleich dem Kehrwert der Informationsfunktion ist.

Als wirksamste Schätzfunktionen kommen zudem auch nur solche Stichprobenfunktionen in Frage, welche die in den Daten enthaltene statistische Information über den zu schätzenden Parameter voll ausschöpfen, so dass die (bedingte) Likelihood der Stichprobendaten bei gegebenem Schätzwert $\hat{\vartheta} = u(x_1, \dots, x_n)$ nicht mehr von dem zu schätzenden Parameter ϑ abhängt: **Suffizienz**

$$L\{x_1, \dots, x_n | \hat{\vartheta}\} = \frac{L\{x_1, \dots, x_n\}}{L\{\hat{\vartheta}\}} = c$$

Schätzfunktionen, welche diese Eigenschaft besitzen, heißen erschöpfend oder suffizient.

Konsistenz

Darüber hinaus wird man fordern, dass eine gute Schätzfunktion den zu schätzenden Parameter mit wachsendem Stichprobenumfang $n \rightarrow \infty$ beliebig genau zu schätzen vermag, so dass für jede beliebig kleine Zahl $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \text{prob}\{|\hat{\vartheta} - \vartheta| < \varepsilon\} = 1$$

Solche Schätzfunktionen bezeichnet man als konsistent.¹

Eigenschaften der ML-Methode

Eigenschaften der Maximum-Likelihood-Methode

- Existiert eine erschöpfende Schätzfunktion U für den zu schätzenden Parameter ϑ , so ist die Lösung der Gleichung [16] eine Funktion von u allein², und die ML-Schätzung $\hat{\vartheta}$ ist daher ebenfalls erschöpfend.
- Existiert eine wirksamste Schätzfunktion U für den zu schätzenden Parameter ϑ , so ist u fast überall in dem Bereich, in dem U eine positive Wahrscheinlichkeitsdichte besitzt, eine Lösung der Gleichung [16]³ und kann daher fast immer mittels ML-Methode gefunden werden.
- Sei U eine ML-Schätzfunktion eines unbekanntem Parameters ϑ , so ist sie zumindest asymptotisch normalverteilt, erwartungstreu und wirksamst.⁴

Intervallschätzung

Auch Schätzfunktionen, welche die genannten Eigenschaften haben, bleiben stets Zufallsvariablen, die eine gewisse Fehlerstreuung besitzen, so dass man selbst in noch so großen Stichproben nie sicher sein kann, wie gut sie den zu schätzenden Parameter tatsächlich approximieren. Wenn die Verteilung der Schätzfunktion bekannt ist, kann man jedoch zumindest ein Konfidenzintervall

1. Als Spezialfall der Konsistenz beschreibt das Gesetz der großen Zahlen (vgl. Kapitel 2.6.1.2) die Konsistenz der relativen Häufigkeit als Schätzfunktion für den Parameter p der Bernoulli-Verteilung.
 2. Vgl. Fisz (1970, 565), Satz 13.7.1.
 3. Vgl. Fisz (1970, 565), Satz 13.7.2.
 4. Die asymptotischen Eigenschaften der ML-Schätzfunktionen wurden u.a. von Hotelling (1930), Doob (1934), Dugué (1937), Wald (1943) und Wolfowitz (1949) untersucht. Zum Beweis des hier zitierten Satzes siehe auch die Monographie von Cramér (1946).

$$\text{KONF}\{u_1 \leq \vartheta \leq u_2\} = \gamma$$

berechnen, innerhalb dessen der zu schätzende Parameter mit einer vorgegebenen Sicherheit (γ) liegt. Der Begriff des Konfidenzintervalls geht auf Neyman (1935) zurück. Die Endpunkte des Intervalls (u_1, u_2) werden als Konfidenzgrenzen bezeichnet. Die Zahl γ heißt die zugehörige Konfidenzzahl. Praktisch wählt man γ gleich 95% oder 99%, manchmal auch 99.9%.

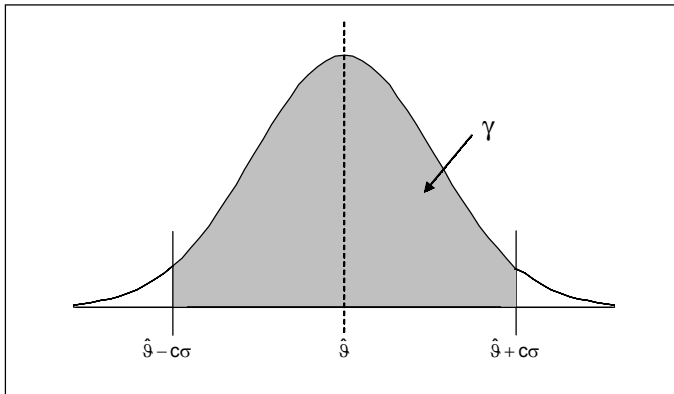


Abb. 2.6.10: Konfidenzintervall für den unbekannten Parameter ϑ

Folgt die Schätzfunktion U einer Normalverteilung mit dem Erwartungswert $E(U) = \vartheta$ und der Standardabweichung $\sigma(U) = \sigma$, so errechnen sich die Konfidenzgrenzen nach

$$u_1 = \hat{\vartheta} - c\sigma \quad \text{und} \quad u_2 = \hat{\vartheta} + c\sigma$$

(vgl. Abb. 2.6.10), wobei die Werte für c der Tabelle für die $N(0,1)$ (= Standardnormalverteilung) entnommen werden können und je nach Konfidenzzahl $c = 1.96$ ($\gamma = 95\%$), $c = 2.576$ ($\gamma = 99\%$) oder $c = 3.291$ ($\gamma = 99.9\%$) betragen.

2.6.3.2 Das Testen von Hypothesen

Im induktiv-statistischen Erklärungsmodell sind unsere Theorien und Hypothesen nicht mehr unmittelbar als Aussagen über Beobachtungen formuliert, sondern als Aussagen über die Wahrscheinlichkeitsverteilung einer wohldefinierten Grundgesamtheit möglicher Beobachtungen. Dies hat zur Folge, dass auch die Bestätigung oder Falsifikation der Hypothesen nur noch mit einer gewissen Irrtumswahrscheinlichkeit erfolgen kann.

Ein einführendes
Beispiel

Betrachten wir dazu ein einfaches Experiment zur Frustrations-Aggressions-Hypothese. Um zu überprüfen, ob Frustration tatsächlich zu einem Anstieg der Wahrscheinlichkeit verletzenden Verhaltens führt, konstruieren wir einen 2-Gruppen-Versuchsplan mit unabhängigen Stichproben:

- Frustrationsbedingung: Die Vpn der Versuchsgruppe sollen einen Konzentrationsleistungstest bearbeiten. Wenn es ihnen gelingt, eine bestimmte Leistung zu erbringen, bekommen sie 50.- € Erfolgshonorar. In dem Raum ist eine zweite Person anwesend. Vermeintlich eine andere Vp, die denselben Test bearbeiten soll, tatsächlich jedoch ein instruierter VL-Gehilfe, der die Vp dauernd stört, so dass sie die geforderte Leistung nicht erbringen kann und kein Honorar erhält (UV: $F(x)$).
- Kontrollbedingung: Die Vpn der Kontrollgruppe bearbeiten denselben Konzentrationsleistungstest. Auch ihnen wird ein Erfolgshonorar von 50.- € versprochen. Der VL-Gehilfe bearbeitet seinen Test und verhält sich dabei ruhig, so dass die Vp nicht gestört wird, die geforderte Leistung erbringen kann und das versprochene Honorar erhält (UV: $\neg F(x)$).
- Anschließend sitzen Vp und VL-Gehilfe in einem anderen Raum des Forschungslaboratoriums zusammen, wo ihnen eine Tasse Kaffee angeboten wird, während sie noch auf irgendetwas warten müssen. Der VL-Gehilfe beginnt ein Gespräch über den Test, den die beiden soeben bearbeitet haben. Durch eine Einwegscheibe beobachtet der VL, ob die Vp in diesem Gespräch verletzende Äußerungen (= verbale Gewalt) gegenüber dem VL-Gehilfen macht (AV: $G(x)$ oder $\neg G(x)$).

Empirische Prognose Aufgrund der Frustrations-Aggressions-Hypothese würden wir vermuten, dass die Wahrscheinlichkeit verletzender Äußerungen in der Versuchsgruppe größer ist als in der Kontrollgruppe:

$$\text{prob}\{G(x)|F(x)\} > \text{prob}\{G(x)|\neg F(x)\}.$$

Wenn diese Hypothese stimmt, dann werden sich in der VG mehr Vpn finden, die im zweiten Teil des Experimentes verletzende Äußerungen gegenüber dem VL-Gehilfen machen. Das Ergebnis könnte z.B. so aussehen wie in Tab. 2.6.3.

Tab. 2.6.3: Ergebnisse eines Experimentes zur Frustrations-Aggressions-Hypothese: absolute Häufigkeiten

	VG: F	KG: $\neg F$
Verbale Gewalt: G	5	1
Keine verbale Gewalt: $\neg G$	3	9
Gesamt	8	10

Die Häufigkeiten in Tab. 2.6.3 lassen sich aber schwer vergleichen, weil die Stichproben unterschiedlich groß sind. Wir verwenden daher besser relative Häufigkeiten, die wir durch Division mit dem Stichprobenumfang (Spaltensumme) erhalten. Diese sind zugleich die empirischen Schätzungen der in der statistischen Gesetzesaussage vorkommenden Wahrscheinlichkeiten (vgl. Tab. 2.6.4).

Ergebnisdarstellung

	VG: F	KG: \neg F
Verbale Gewalt: G	62.5%	10.0%
Keine verbale Gewalt: \neg G	37.5%	90.0%
Gesamt	100.0%	100.0%

Tab. 2.6.4: Ergebnisse eines Experimentes zur Frustrations-Aggressions-Hypothese: relative Häufigkeiten

Der Unterschied zwischen den beiden Gruppen sieht recht dramatisch aus (vgl. Abb. 2.6.11) und scheint die Frustrations-Aggressions-Hypothese zu bestätigen. Andererseits müssen wir uns darüber klar sein, dass wir doch nur relativ kleine Stichproben von Beobachtungen haben: 8 Vpn in der VG und 10 in der KG. Es könnte also durchaus sein, dass dieses Ergebnis rein zufällig zustande gekommen ist und dass die Frustrationsbedingung in Wirklichkeit keinerlei Einfluss auf die Wahrscheinlichkeit verbal verletzenden Verhaltens während der zweiten Phase des Experimentes hat.

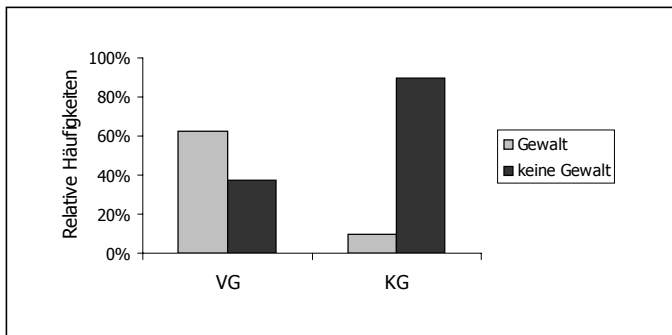


Abb. 2.6.11: Ergebnisse eines Experimentes zur Frustrations-Aggressions-Hypothese: relative Häufigkeiten

Diese Annahme, wonach sich VG und KG nicht voneinander unterscheiden, bezeichnet man in der Statistik als Nullhypothese (H_0). Sie lautet in unserem Beispiel:

Nullhypothese und Alternativhypothese

$$H_0: \text{prob}\{G(x)|F(x)\} = \text{prob}\{G(x)|\neg F(x)\}.$$

Der Nullhypothese wird die Alternativhypothese (H_1) gegenübergestellt, nach der zwischen VG und KG sehr wohl ein Unterschied besteht; und zwar entweder in der prognostizierten oder in der entgegengesetzten Richtung (zweiseitige Fragestellung):

$$H_1: \text{prob}\{G(x)|F(x)\} \neq \text{prob}\{G(x)|\neg F(x)\}.$$

Statistische
Signifikanz

Damit die Frustrations-Aggressions-Hypothese durch den Ausgang unseres Experimentes bestätigt wird, ist zweierlei erforderlich:

1. Die Unterschiede zwischen VG und KG müssen in die prognostizierte Richtung gehen.
2. Wir müssen eine Entscheidung zwischen H_0 und H_1 herbeiführen.

Dabei wird H_0 dann zugunsten von H_1 verworfen, wenn der Unterschied zwischen VG und KG so deutlich ausfällt, dass rein zufällige Unterschiede dieses Ausmaßes äußerst unwahrscheinlich sind. Man sagt dann, der Unterschied ist statistisch signifikant. Ist dies nicht der Fall, so wird H_0 beibehalten.

Genauer gesagt, wird H_0 dann verworfen, wenn die Wahrscheinlichkeit so großer oder noch größerer Unterschiede bei Geltung der Nullhypothese kleiner ist als eine vorgegebene Irrtumswahrscheinlichkeit α . Diese ist bereits vor Kenntnis der Daten vom VL festzulegen und wird als Signifikanzniveau bezeichnet.

Zur Geschichte des
Hypothesentestens

Der erste Test einer Nullhypothese wurde bereits 1710 von John Arbuthnot durchgeführt und bediente sich einer Argumentationsweise, die dem heute gebräuchlichen Vorgehen überraschend ähnlich ist (vgl. Gigerenzer & Murray, 1987). Der erste moderne Signifikanztest war die von Karl Pearson (1900) entwickelte χ^2 -Methode.

Die Logik des Nullhypothesentestens, wie sie in der Psychologie heute Standard ist, stellt eine Art Hybrid zweier divergierender Ansätze dar, die auf Neyman & Pearson (1928a,b; Neyman, 1938) bzw. Fisher (1935) zurückgehen und die von den Autoren erbittert gegeneinander verteidigt wurden (vgl. Gigerenzer & Murray, 1987).

Arten von Fehlern

Ein zentrales Problem des Hypothesentestens besteht darin, dass man die Nullhypothese nie beweisen, sondern nur (mit der vorgegebenen Fehlerwahrscheinlichkeit α) zugunsten der Alternativ-

hypothese verwerfen kann. Dabei sind zwei Arten von Fehlentscheidungen möglich (vgl. Tab. 2.6.5).

Empirische Wirklichkeit	Ergebnis der statistischen Hypothesenprüfung	
	H_0 verworfen	H_0 beibehalten
Tatsächlich ist H_0 richtig	Fehler 1. Art: H_0 wird fälschlich verworfen	
Tatsächlich ist H_1 richtig		Fehler 2. Art: H_0 wird fälschlich beibehalten

Tab. 2.6.5: Arten von Fehlentscheidungen beim Hypothesentesten

Vom VL festgesetzt werden kann aber nur die Wahrscheinlichkeit eines Fehlers 1. Art (das Signifikanzniveau α), denn nur unter Annahme der Nullhypothese lässt sich die Wahrscheinlichkeitsverteilung der empirisch gefundenen Daten exakt bestimmen. Die Wahrscheinlichkeit eines Fehlers 2. Art (β) lässt sich dagegen nicht genau bestimmen. Von ihr wissen wir nur, dass β desto größer wird, je kleiner wir α wählen (vgl. Abb. 2.6.12).

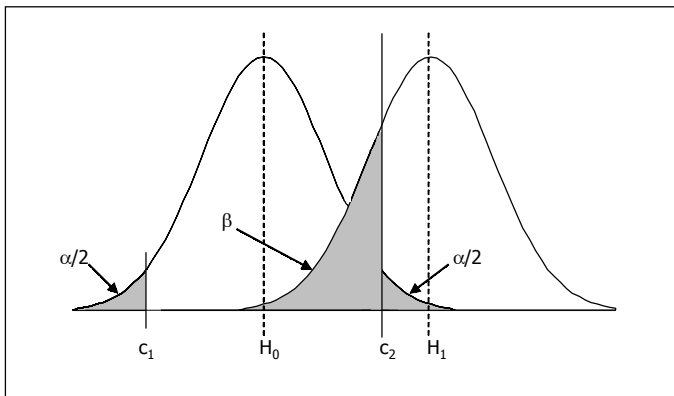


Abb. 2.6.12: Fehlerwahrscheinlichkeiten beim Hypothesentesten

- Gemäß der Forderung nach möglichst strenger Hypothesenprüfung werden wir den α -Fehler daher möglichst klein halten, wenn die zu prüfende Sachhypothese – wie in unserem Beispiel – mit H_1 übereinstimmt ($\alpha=1\%$ oder kleiner). Dadurch verringern wir die Gefahr, dass die Sachhypothese fälschlich bestätigt wird (vgl. Tab. 2.6.6).
- Stimmt die Sachhypothese dagegen mit H_0 überein, so müs-

Strenge
Hypothesenprüfung

sen wir Vorsorge treffen, dass H_0 nicht fälschlich beibehalten wird. Daher werden wir α eher groß wählen ($\alpha=5\%$) (vgl. Tab. 2.6.6).

Tab. 2.6.6: Kritische Hypothesenprüfung

Kritische Hypothesenprüfung	
Der aus der Theorie hergeleiteten Sachhypothese entspricht die Nullhypothese H_0	Der aus der Theorie hergeleiteten Sachhypothese entspricht die Alternativhypothese H_1
α = groß: Schon relativ kleine Abweichungen von der Erwartung unter H_0 gelten als Falsifikation der Sachhypothese	α = klein: Erst relativ starke Abweichungen von der Erwartung unter H_0 gelten als Bestätigung der Sachhypothese

Bewährung oder Falsifikation der Sachhypothese

- Wird die Nullhypothese zugunsten der Alternativhypothese verworfen und die Gruppenunterschiede entsprechen unserer Prognose (mehr verbale Gewalt nach Frustration, weniger verbale Gewalt ohne Frustration), dann hat sich unsere Hypothese bewährt: Frustration lässt verbal verletzendes Verhalten wahrscheinlicher werden.
- Wird die Nullhypothese zugunsten der Alternativhypothese verworfen und die Gruppenunterschiede widersprechen unserer Prognose (weniger verbale Gewalt nach Frustration, mehr verbale Gewalt ohne Frustration), dann ist unsere Hypothese falsifiziert: Frustration lässt verbal verletzendes Verhalten weniger wahrscheinlich werden.

Wird die Nullhypothese dagegen beibehalten, so ist die Hypothese weder bestätigt noch falsifiziert: In diesem Fall reichen unsere Daten einfach nicht aus, um eine Entscheidung über die Hypothese herbeizuführen (vgl. Tab. 2.6.7).

Tab. 2.6.7: Signifikanz, Bewährung und Falsifikation der Sachhypothese

H_0	Gruppenunterschiede	
	Entsprechen der Prognose: Mehr Gewalt nach Frustration	Widersprechen der Prognose: Weniger Gewalt nach Frustration
Verworfen	<i>Bewährung</i> der Hypothese	<i>Falsifikation</i> der Hypothese
Beibehalten	<i>Weder Bewährung noch Falsifikation</i>	

Hat Frustration tatsächlich keinerlei Einfluss auf die Wahrscheinlichkeit verletzenden Verhaltens, dann werden wir dies mit einem einzelnen Experiment nicht nachweisen können.

Aber über viele Experimente hinweg werden wir viele Experimente finden, die zu keinen signifikanten Ergebnissen führen (d.h. bei denen H_0 nicht verworfen werden kann), und einige Experimente, in denen die Nullhypothese (irrtümlich) verworfen wurde (vgl. Abb. 2.6.13).

Dass diese Verwerfungen der Nullhypothese irrtümlich zustande gekommen sind, wird man an zweierlei bemerken:

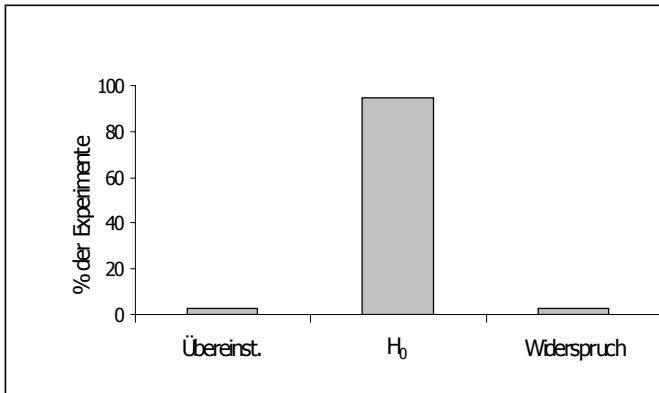


Abb. 2.6.13: Ergebnisse vieler Experimente bei Geltung der Nullhypothese

- Erstens werden die signifikanten Gruppenunterschiede teilweise mit unserer Hypothese übereinstimmen (mehr verletzendes Verhalten nach Frustration), und teilweise werden sie ihr widersprechen (weniger verletzendes nach Frustration).
- Zweitens wird die relative Häufigkeit, mit denen die Experimente zu signifikanten Ergebnissen führen, in etwa der (vom VL festgesetzten) Irrtumswahrscheinlichkeit α entsprechen, mit der die Nullhypothese fälschlich verworfen wird (z.B. 5%).

Aus diesen Gründen kann man im induktiv-statistischen Erklärungsmodell über die Geltung einer (probabilistischen) Gesetzesaussage nicht mittels eines einzelnen Entscheidungsexperimentes befinden, sondern man braucht lange Versuchsserien von wiederholten Experimenten.

2.6.3.3 Zweiseitige und einseitige Fragestellungen

In ihrer allgemeinsten Form besagt die Nullhypothese, dass sich zwei Wahrscheinlichkeitsverteilungen bezüglich eines bestimmten Parameters (ϑ) nicht voneinander unterscheiden

Der Gegensatz von Nullhypothese und Alternativhypothese

$$H_0: \vartheta_1 = \vartheta_2$$

Bei den beiden Verteilungen kann es sich z.B. um den Vergleich zwischen VG und KG oder auch um den Vergleich einer empirischen Verteilung mit einer theoretischen Verteilung handeln.

Um H_0 testen zu können, muss ihr eine Alternativhypothese gegenübergestellt werden, die zu H_0 in kontradiktorischem Gegensatz steht. Nur wenn diese Voraussetzung erfüllt ist, kann eine Entweder-oder-Entscheidung zwischen H_0 und H_1 herbeigeführt werden.

Zweiseitige
Fragestellung

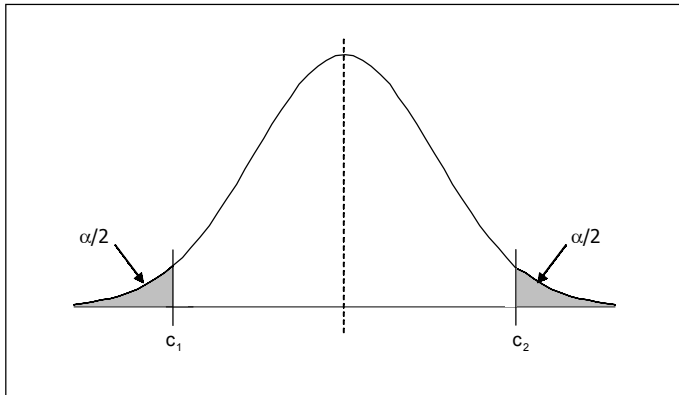
Im einfachsten Fall gewinnen wir H_1 durch Negation von H_0 und stellen das folgende Hypothesenpaar einander gegenüber:

$$H_0: \vartheta_1 = \vartheta_2$$

$$H_1: \vartheta_1 \neq \vartheta_2$$

Man spricht dann von einer zweiseitigen Fragestellung (weil H_1 sowohl $\vartheta_1 > \vartheta_2$ als auch $\vartheta_1 < \vartheta_2$ einschließt, d.h. die Abweichung von H_0 kann in beide Richtungen gehen)(vgl. Abb. 2.6.14).

Abb. 2.6.14: Annahme- und Verwerfungsbereich der Nullhypothese bei zweiseitiger Fragestellung. Die Nullhypothese wird verworfen, wenn die Prüfgröße einen Wert $\leq c_1$ oder einen Wert $\geq c_2$ annimmt



Einseitige
Fragestellung

Mitunter verwendet man in der Statistik aber auch einseitige Fragestellungen, bei denen der Nullhypothese eine gerichtete Alternativhypothese gegenübergestellt wird, so dass wir o.B.d.A. ein Hypothesenpaar der Form

$$H_0: \vartheta_1 = \vartheta_2$$

$$H_1: \vartheta_1 > \vartheta_2$$

erhalten (vgl. Abb. 2.6.15).

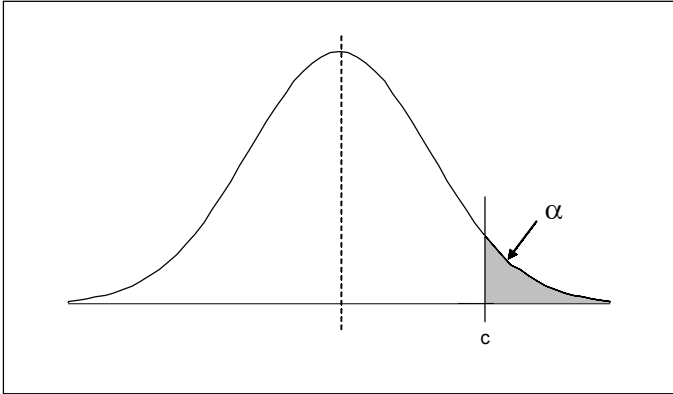


Abb. 2.6.15: Annahme- und Verwerfungsbereich der Nullhypothese bei einseitiger Fragestellung. Die Nullhypothese wird verworfen, wenn die Prüfgröße einen Wert $\geq c$ annimmt

Für die Formulierung einer einseitigen Fragestellung reicht es freilich nicht aus, dass die zu prüfende Sachhypothese einen gerichteten Unterschied ($\vartheta_1 > \vartheta_2$) prognostiziert. Damit ein kontradiktorischer Gegensatz zwischen H_0 und H_1 besteht, muss die entgegengesetzte Abweichung zwischen den beiden Parametern ($\vartheta_1 < \vartheta_2$) – unabhängig von der zu prüfenden Sachhypothese – definitiv ausgeschlossen werden können: entweder strukturell (d.h. sachlogisch) oder empirisch (d.h. aufgrund einer bereits bewährten empirischen Theorie).¹ Dass dies nicht beachtet wird, stellt – nicht nur in der Psychologie – eine der verbreitetsten Fehlerquellen empirischer Forschung dar.

Voraussetzungen für einseitige Fragestellungen

Verwendet man fälschlich eine einseitige Fragestellung, ohne der H_1 entgegengesetzte Abweichungen zwischen den beiden Verteilungen definitiv ausschließen zu können, so hat dies weitreichende Konsequenzen:

1. Die Alternativhypothese ($H_1: \vartheta_1 > \vartheta_2$) wird dann nicht mehr gegen die Nullhypothese ($H_0: \vartheta_1 = \vartheta_2$) getestet, sondern gegen $H_0^*: \vartheta_1 \leq \vartheta_2$. Die Beibehaltung der Nullhypothese kann folglich nicht mehr als Null-Unterschied zwischen den beiden Verteilungen interpretiert werden, und es kann nicht ausgeschlossen werden, dass sich die Verteilungen sehr wohl signifikant voneinander unterscheiden – allerdings in entgegengesetzter Richtung als von H_1 prognostiziert.

1. Aus diesem Grund haben wir auch in dem oben dargestellten Experiment zur Frustrations-Aggressions-Hypothese eine zweiseitige Fragestellung gewählt, obwohl die Sachhypothese eine gerichtete Abweichung zwischen VG und KG vorhersagt.

2. Die Wahrscheinlichkeit, dass H_0 fälschlich verworfen wird, verdoppelt sich gegenüber dem nominell festgesetzten Signifikanzniveau α .

Fazit: Die Verwendung einer einseitigen Fragestellung ist nur dann zulässig, wenn die entgegengesetzte Hypothese definitiv ausgeschlossen werden kann.

Ein Beispiel

Wohin es führen kann, wenn man fälschlich eine einseitige Fragestellung wählt, obwohl die entgegengesetzte Hypothese nicht definitiv ausgeschlossen werden kann, zeigt das folgende Beispiel einer Untersuchung des (teilweise auch in der Psychologie) verbreiteten Vorurteils, wonach Schwarze weniger intelligent sein sollen als Weiße.

Dabei kann *keine* der drei Möglichkeiten a priori ausgeschlossen werden: (1) Schwarze sind weniger intelligent, (2) Schwarze sind gleich intelligent, oder (3) Schwarze sind intelligenter als Weiße.

Wählt man fälschlich eine einseitige Fragestellung und testet damit die Alternativhypothese

H_1 : "Schwarze sind weniger intelligent",

so kann das Ergebnis der Untersuchung

1. niemals zu dem Ergebnis führen, dass Schwarze intelligenter sind – auch dann nicht, wenn dies der Wirklichkeit entspricht –, und
2. wird aufgrund der Verdoppelung des tatsächlichen Signifikanzniveaus gegenüber dem nominellen Signifikanzniveau die intellektuelle Überlegenheit der Weißen – zumindest in einigen Fällen – auch dann fälschlich bewiesen werden, wenn in Wirklichkeit die Nullhypothese zutrifft und es keine Intelligenzunterschiede zwischen Schwarz und Weiß gibt.

Die fälschliche Verwendung einer einseitigen Fragestellung führt so am Ende zur (pseudo-)wissenschaftlichen Untermauerung des Vorurteils. Der Vollständigkeit halber sei aber angemerkt, dass Untersuchungen der beschriebenen Art selbst dann nicht aussagekräftig sind, wenn die statistischen Hypothesen korrekt formuliert und das Vorurteil mittels zweiseitiger Fragestellung überprüft wird. Denn um Intelligenzunterschiede zwischen Schwarz und Weiß nachzuweisen (egal in welcher Richtung), müssten wir repräsentative Stichproben von Schwarzen und Weißen miteinander

der vergleichen. Gleichzeitig müssten wir alle möglichen Störvariablen wie Sozialstatus, Bildungschancen usw. kontrollieren, welche die Testergebnisse der Probanden sonst noch beeinflussen könnten. Wegen der sozialen Ungleichheit von Schwarz und Weiß schließen sich Repräsentativität und Vergleichbarkeit bezüglich Sozialstatus und Bildungschancen etc. jedoch gegenseitig aus.

2.6.4 Zufallsauswahl und Zufallsexperiment

Wahrscheinlichkeitsaussagen setzen die Zufälligkeit der Daten voraus, über welche die Aussagen getroffen werden. Dies gilt insbesondere auch für den inferenzstatistischen Schluss von der Stichprobe auf die Grundgesamtheit. Inferenzstatistische Schlüsse sind daher nur dann gerechtfertigt, wenn es sich bei der Stichprobe um eine zufällige Stichprobe aus der Grundgesamtheit aller möglichen Beobachtungen handelt.

Ein in der Psychologie verbreitetes Missverständnis besteht darin, dass diese Grundgesamtheit, die man in der Statistik auch als Population bezeichnet, mit der Personenpopulation identifiziert wird, für welche die Geltung der zu prüfenden Hypothese postuliert wird. Statistische Grundgesamtheit und Personenpopulation sind jedoch zwei verschiedene Dinge, die man auseinanderhalten muss.

Ein verbreiteter Irrtum

Eine Identität von Grundgesamtheit und Personenpopulation besteht allerdings in der Umfrageforschung (Survey-Forschung); z.B. wenn wir mittels Fragebogenerhebung feststellen wollen, wie groß der gegenwärtige Wähleranteil einer bestimmten Partei ist, oder auch z.B. im Milgram-Experiment, soweit es darin um die Frage geht, wie groß der Anteil der Probanden ist, die aus Gehorsam gegenüber dem VL und der Institution ‚Wissenschaft‘ vor massiver Gewaltausübung nicht zurückschrecken.

Umfrageforschung

Um solche Aussagen machen zu können, benötigen wir eine repräsentative Stichprobe von V_{pn} , d.h. eine Personenstichprobe, die in jeder Hinsicht die Verteilungsmerkmale der Bevölkerung korrekt widerspiegelt.

Wie beim Ziehen aus einer Urne haben wir es hier also mit einem Problem vom Typ der Zufallsauswahl zu tun, und wegen des Gesetzes der großen Zahlen wissen wir, dass die Repräsentativität

Zufallsstichproben und anfallende Stichproben

zufällig gezogener Stichproben mit wachsendem Stichprobenumfang zunimmt. In der Psychologie erheben wir aber oft nur relativ kleine Stichproben, und in der Praxis ist auch die Zufallsauswahl der Vpn oft schwer zu realisieren. Das liegt u.a. an hohen Kosten, mangelnder Verfügbarkeit einer kompletten Liste der Elemente der Zielpopulation, mangelnder Teilnahmebereitschaft einiger potentieller Vpn usw. An Stelle einer Zufallsstichprobe von Vpn hat man es daher meist lediglich mit einer anfallenden Stichprobe zu tun (z.B. mit jenen Psychologiestudenten, die auf einen Ausgang reagieren und sich freiwillig für das Experiment gemeldet haben).

Geschichtete
Stichproben

Um die Repräsentativität einer Stichprobe schon bei relativ kleinem Stichprobenumfang und bei mangelnder Zufallsauswahl zu verbessern, kann man aber geschichtete Stichproben bilden.

Solche geschichteten Stichproben werden so konstruiert, dass sie zumindest die Verteilung bestimmter, als eventuelle Störvariable identifizierter Variablen korrekt widerspiegeln: Wenn z.B. in der Population 55% der Vpn männlich sind, so enthält eine (hinsichtlich des Geschlechtes) geschichtete Stichprobe den gleichen Prozentsatz.

Typen von
Personenstichproben

Je nach Art der Stichprobenziehung lassen sich die folgenden Typen von Personenstichproben unterscheiden (vgl. Tab. 2.6.8):

Tab. 2.6.8: Typen von Personenstichproben

	Mit Zufallsauswahl	Ohne Zufallsauswahl
Ohne Schichtung	Reine Zufallsstichprobe	Anfallende Stichprobe
Mit Schichtung	Geschichtete Zufallsstichprobe	Quotenstichprobe

Reine
Zufallsstichprobe

Bei der reinen Zufallsstichprobe werden die Probanden zufällig aus der Population gezogen.

Geschichtete
Zufallsstichprobe

Bei der geschichteten Zufallsstichprobe wird die Population nach verschiedenen Schichtungsmerkmalen (z.B. Geschlecht, Alter und Schulbildung) in Schichten eingeteilt. Per Zufall werden aus jeder Schicht so viele Probanden gezogen, dass die Größe der Stichprobenschichten zur Größe der entsprechenden Populationsschichten proportional ist.

Anfallende
Stichprobe

Bei der anfallenden Stichprobe werden all jene Probanden in die Stichprobe aufgenommen, die gerade anfallen.

Wie bei der geschichteten Zufallsstichprobe wird die Population auch bei der Quotenstichprobe in Schichten eingeteilt und die Größe der Stichprobenschichten zur Größe der entsprechenden Populationsschichten proportional gehalten. Die Probanden einer Stichprobenschicht werden aus der entsprechenden Populationschicht jedoch nicht zufällig gezogen, sondern man nimmt die gerade anfallenden Probanden, welche die Schichtungsmerkmale erfüllen.

Quotenstichprobe

In der experimentellen Forschung spielt die Repräsentativität der Personenstichprobe nur eine untergeordnete Rolle. Wie beim Werfen eines Würfels oder beim Drehen eines Glücksrades haben wir es hier nämlich nicht mit einem Problem vom Typ der Zufallsauswahl, sondern mit einem Problem vom Typ des Zufallsexperimentes zu tun. Die statistische Grundgesamtheit ist durch die Menge aller denkbaren Ausführungen des Experimentes definiert und kann nicht mit der Personenpopulation gleichgesetzt werden.

Experimentelle
Forschung

Im Gegenteil: Man wird auf die Repräsentativität der Personenstichprobe ggf. sogar gezielt verzichten. Wenn wir z.B. ein Experiment durchführen, das hohe intellektuelle Anforderungen an die Vpn stellt, werden wir besser daran tun, das Experiment an einer Stichprobe von intellektuell hinreichend begabten Vpn durchzuführen und nicht an einer repräsentativen Personenstichprobe, die das ganze Spektrum bis hin zu schwerster intellektueller Minderbegabung umfasst.

Repräsentativitäts-
mängel und
Bewährungsgrad der
Hypothese

Weichen wir von der Repräsentativität der Personenstichprobe ab, so geht dies allerdings auf Kosten des Bewährungsgrades der untersuchten Sachhypothese, weil wir die Hypothese dann nur an einem bestimmten Ausschnitt aus der Personenpopulation getestet haben.

Zur Kontrolle allfälliger Störvariablen, die von den Probanden mit ins Experiment gebracht werden (z.B. Begabung, Motivation etc.), ist es in der experimentellen Forschung allerdings erforderlich, dass die anfallenden Vpn zufällig auf die experimentellen Gruppen verteilt werden.

Die Randomisierung
von Störvariablen

Da sich die zu prüfende Hypothese auf die Grundgesamtheit aller möglichen Ausführungen des Experimentes bezieht und nicht auf eine bestimmte Personenpopulation, reicht dies im Falle des o.g. stochastischen Modells sowie auch im Falle des Fehlermodells mit

fester Effektgröße (fixed effect model) völlig aus, und die Verwendung einer anfallenden Stichprobe geht

- weder auf Kosten der internen Validität des Experimentes
- noch auf Kosten der Korrektheit des inferenzstatistischen Schlusses.

Gewisse Einschränkungen gibt es dagegen im Falle des Fehlermodells mit zufälliger Effektgröße (random effect model), wie wir wieder am Beispiel des Weber'schen Gesetzes erläutern können:

Bei Annahme einer zufälligen Effektgröße unterscheiden sich die Vpn in ihrer (relativen) Unterschiedsschwelle (c_v), und das Weber'sche Gesetz beschreibt die mittleren Unterschiedsschwellen in der Personenpopulation:

$$\Delta S = \mu_c \times S.$$

Um diese mittleren Unterschiedsschwellen korrekt bestimmen zu können, ist dann (ähnlich wie bei den o.g. Prozentangaben aus dem Milgram-Experiment) sehr wohl die Repräsentativität der Personenstichprobe erforderlich.

Indem die Geltung des Weber'schen Gesetzes aber für jede einzelne Vp behauptet wird

$$\Delta_v = c_v \times S,$$

gilt andererseits die Beziehung $E(D) = E(C) \times S$ in jeder – wie auch immer selektierten – Personenpopulation, und die experimentelle Geltungsprüfung des Gesetzes ist daher auch hier nicht an die Voraussetzung der Repräsentativität der Personenstichprobe gebunden.

- Die *Konstanz* der (mittleren) relativen Unterschiedsschwelle kann auch anhand nicht repräsentativer Stichproben *überprüft* werden.
- Um zu *schätzen, wie groß* die mittleren Unterschiedsschwellen $\Delta S = E(D)$ und $\mu_c = E(C)$ in der Personenpopulation sind, sind dagegen repräsentative Stichproben erforderlich.

Kontrollfragen

- Wie lautet das induktiv-statistische Erklärungsmodell? Geben Sie ein Beispiel!
- Worin besteht der Unterschied zum deduktiv-nomologischen Erklärungsmodell?

- Welchen Vorteil bietet es (insbesondere in der Psychologie)?
- Welche Nachteile muss man dafür in Kauf nehmen?
- Wie lauten die Axiome von Kolmogoroff?
- Warum sind die Axiome von Kolmogoroff nicht ausreichend, um die Bedeutung des Wahrscheinlichkeitsbegriff zu klären?
- Warum benötigen wir in der Psychologie ein Modell des Wahrscheinlichkeitsbegriffs?
- Erläutern Sie die Termini ‚zufälliges Ereignis‘ und ‚Zufallsexperiment‘!
- Erläutern Sie die Konstruktionsprinzipien für
 - a) diskrete,
 - b) kontinuierliche Zufallsgeneratoren!
- Wozu dient der Wahrscheinlichkeitsbegriff? Was soll die Wahrscheinlichkeit beschreiben?
- Leiten Sie den Laplace’schen Wahrscheinlichkeitsbegriff aus den Konstruktionsprinzipien für diskrete Zufallsgeneratoren her und begründen Sie, dass er tatsächlich die gewünschte Quantifizierung der Kontingenz leistet!
- Begründen Sie, warum beim Glücksrad die Wahrscheinlichkeit $\text{prob}\{a \leq X \leq b\} = b - a$ ist!
- Erläutern Sie, warum
 - a) das Ziehen aus einer Urne ohne Zurücklegen
 - b) das Werfen eines unregelmäßigen Würfels als Zufallsexperiment betrachtet werden kann!
- Was besagt das Gesetz der großen Zahlen?
- Wie lauten Erwartungswert und Varianz der relativen Häufigkeit?
- Wie lautet die Ungleichung von Tschebyscheff?
- Zeigen Sie, dass die Ungleichung von Tschebyscheff auch für diskrete Zufallsvariablen gilt!
- Inwiefern sind Wahrscheinlichkeiten idealisierte relative Häufigkeiten?
- Erläutern Sie am Beispiel des Weber’schen Gesetzes, welche Arten der Formulierung von statistischen Gesetzesaussagen sich unterscheiden lassen und auf welchen Modellannahmen sie beruhen!
- Erläutern Sie den Unterschied zwischen einem Zufallsexperiment und einer Zufallsauswahl!
- Erläutern Sie am Beispiel des Weber’schen Gesetzes den Unterschied zwischen fester Effektgröße und zufälliger Effektgröße!
- Erläutern Sie, inwiefern die Anwendung von Wahrscheinlichkeitsaussagen in der Psychometrie problematisch ist und welche Auswirkungen dies
 - a) in der klassischen Testtheorie,
 - b) in der Latent-Class-Analyse und
 - c) bei den Latent-Trait-Modellen hat!
- Diskutieren Sie die Vor- und Nachteile der CML- und MML-Methode!
- Worin besteht der Unterschied zwischen Statistiken und Parametern?
- Wie unterscheidet sich die Definition der Häufigkeitsfunktion bei diskreten und stetigen Zufallsvariablen?
- Wie unterscheidet sich die Definition der Wahrscheinlichkeitsdichte bei diskreten und stetigen Zufallsvariablen?

- Wie sind Summenhäufigkeitsfunktion und Verteilungsfunktion definiert?
- Welcher Zusammenhang besteht zwischen der Häufigkeitsfunktion und der Summenhäufigkeitsfunktion?
- Welcher Zusammenhang besteht zwischen der Wahrscheinlichkeitsdichte und der Verteilungsfunktion diskreter Zufallsvariablen?
- Welcher Zusammenhang besteht zwischen der Wahrscheinlichkeitsdichte und der Verteilungsfunktion stetiger Zufallsvariablen?
- Warum ist bei stetigen Zufallsvariablen $\text{prob}\{X=x\} = 0$?
- Wie sind
 - a) der Mittelwert,
 - b) die Stichprobenvarianz,
 - c) die Standardabweichung,
 - d) die Kovarianz,
 - e) die Korrelation
 definiert? Welche Parameter entsprechen ihnen, und wie sind diese definiert?
- Warum können die Parameter einer Wahrscheinlichkeitsverteilung aus den jeweiligen Daten immer nur geschätzt werden?
- Worin besteht der Unterschied zwischen impliziten und expliziten Parametern?
- Was versteht man unter den Momenten einer Verteilung?
- Was versteht man unter dem Produktmoment zweier Zufallsvariablen?
- Worin besteht die Momentenmethode?
- Worin besteht die Maximum-Likelihood-Methode?
- Wie ist die Likelihood-Funktion definiert, und was beschreibt sie
 - a) im Falle diskreter Zufallsvariablen,
 - b) im Falle stetiger Zufallsvariablen?
- Wie unterscheiden sich die CML-, MML- und WML-Methode von der hier dargestellten einfachen ML-Methode?
- Wie geht man zur der Berechnung der ML-Schätzfunktionen praktisch vor?
- Welches sind die wichtigsten Eigenschaften von Schätzfunktionen?
- Was versteht man unter einer erwartungstreuen Schätzfunktion?
- Was versteht man unter der Wirksamkeit von Schätzfunktionen?
- Was besagt die Ungleichung von Rao-Cramér?
- Wie ist die Informationsfunktion definiert?
- Was versteht man unter einer wirksamsten Schätzfunktion?
- Welche Voraussetzungen muss eine Schätzfunktion erfüllen, damit sie als wirksamste Schätzfunktion in Betracht kommt?
- Was versteht man unter der Suffizienz von Schätzfunktionen?
- Was versteht man unter der Konsistenz von Schätzfunktionen?
- Was hat das Gesetz der großen Zahlen mit der Konsistenz von Schätzfunktionen zu tun?
- Welche Eigenschaften hat die ML-Methode?
- Erläutern Sie die Termini
 - a) Konfidenzintervall,
 - b) Konfidenzgrenzen und
 - c) Konfidenzzahl!
- Wie kann man das Konfidenzintervall normalverteilter Schätzfunktio-

nen bestimmen?

- Warum kann die Bestätigung oder Falsifikation von Hypothesen im induktiv-statistischen Erklärungsmodell nur mit einer gewissen Irrtumswahrscheinlichkeit erfolgen?
- Welchen Vorteil bietet die Beschreibung der Daten durch relative Häufigkeiten gegenüber der Beschreibung durch absolute Häufigkeiten?
- Erläutern Sie am Beispiel des Vergleichs zweier Stichproben (VG und KG), was man beim statistischen Hypothesentesten unter
 - a) der Nullhypothese,
 - b) der Alternativhypothese,
 - c) dem Fehler 1. Art und
 - d) dem Fehler 2. Art versteht!
- Was versteht man unter statistischer Signifikanz?
- Was versteht man unter dem Signifikanzniveau?
- Welche Beziehung besteht zwischen dem Fehler 1. Art und dem Fehler 2. Art?
- Wann und warum ist das Signifikanzniveau
 - a) möglichst klein bzw.
 - b) eher groß
 zu wählen?
- Bei welchem Ausgang des Experimentes und der statistischen Hypothesenprüfung hat sich die Sachhypothese bewährt? Wann ist sie falsifiziert?
- Welche Ergebnisse erwarten wir (über viele Experimente hinweg), wenn die UV tatsächlich keinerlei Einfluss auf die AV hat?
- Warum kann man im induktiv-statistischen Erklärungsmodell über die Geltung einer Gesetzesaussage nicht mittels eines einzelnen Entscheidungsexperimentes befinden?
- Welche logische Beziehung besteht zwischen Nullhypothese und Alternativhypothese?
- Worin besteht der Unterschied zwischen einseitiger und zweiseitiger Fragestellung?
- Welche Voraussetzungen sind erforderlich, damit eine einseitige Fragestellung zulässig ist?
- Welche Konsequenzen ergeben sich, wenn diese Voraussetzungen verletzt sind?
- Worin besteht der Unterschied zwischen statistischer Grundgesamtheit und (Personen-)Population?
- Was versteht man unter der Repräsentativität von Stichproben?
- Bei welcher Art von Forschung benötigt man in jedem Fall repräsentative Personenstichproben?
- Welche Schwierigkeiten stehen der Erhebung repräsentativer Personenstichproben in der Psychologie entgegen?
- Worin besteht der Unterschied zwischen Zufallsstichproben und anfallenden Stichproben?
- Welche Typen von Personenstichproben lassen sich unterscheiden, und wie sind sie charakterisiert?
- Wie ist die statistische Grundgesamtheit
 - a) in der Umfrageforschung,

- b) in der experimentellen Forschung definiert?
- Erläutern Sie anhand eines Beispiels, warum man in der experimentellen Forschung gelegentlich bewusst auf die Repräsentativität der Personenstichprobe verzichtet!
 - Welche Konsequenzen hat der Verzicht auf Repräsentativität der Personenstichprobe im Experiment?
 - Warum muss man in der experimentellen Forschung darauf achten, dass die Vpn zufällig auf VG und KG verteilt werden?
 - Erläutern Sie anhand eines Beispiels, welche Konsequenzen der Verzicht auf Repräsentativität der Personenstichprobe
 - a) im Falle des stochastischen Modells,
 - b) im Rahmen des Fehlermodells mit fester Effektgröße,
 - c) im Rahmen des Fehlermodells mit zufälliger Effektgröße hat!

Übungsaufgaben

1. Ein zufälliges Ereignis A habe bei einem Experiment die Wahrscheinlichkeit $p = 0.7$. Berechnen Sie
 - a) den Erwartungswert und
 - b) die Varianz der relativen Häufigkeit, mit welcher das Ereignis in einer Versuchsreihe der Länge $n = 50$ eintritt!
2. Geben Sie mittels der Ungleichung von Tschebyscheff eine obere Grenze für die Wahrscheinlichkeit an, mit welcher die relative Häufigkeit in Aufgabe (1) um mehr als 0.2 von p abweicht!
3. Die Fehlerstreuung bei der Schätzung des IQ einer Vp betrage 2.5 IQ-Punkte. Geben Sie ein Intervall an, innerhalb dessen der tatsächliche IQ der Vp mit einer Sicherheit von
 - a) 95%
 - b) 99%
 liegt, wenn die Vp in dem Test 109 IQ-Punkte erzielt hat!

Kapitel 3
Soft Science: Die Herausforderung an die
Sozialwissenschaften

3.1

Objektseitig und subjektseitig definierte Variablen

Anders als z.B. die Theorien der Physik enthalten psychologische Theorien in der Regel nicht nur objektseitig, sondern auch subjektseitig definierte Variablen (vgl. Tab. 3.1.1).

	Beobachtungssprachlich beschrieben	Theoriesprachlich beschrieben
Objektseitig definiert	Physikalische Stimulusqualitäten (materielle Stimuluseigenschaften)	Funktional gedeutete Stimuli (Semantik der Stimuli)
	Bloßes Verhalten	Funktionales Verhalten
Subjektseitig definiert	Je eigene Wahrnehmungen und Gefühle	Individuelle Sinngehalte

Tab. 3.1.1: Objektseitig und subjektseitig definierte Sachverhalte

Objektseitig definiert sind jedenfalls die von außen beobachtbaren Sachverhalte; z.B. beobachtbare Stimuli wie Lichtreiz von der und der Helligkeit, Wellenlänge, Lokalisierung und Dauer oder beobachtbares Verhalten wie z.B. laufen. Diese Variablen werden beobachtungssprachlich beschrieben.

Objektseitig und beobachtungssprachlich

Subjektseitig definiert sind jedenfalls die Sinngehalte, die durch deutungssprachliche Termini (und damit theoriesprachlich) beschrieben werden und die mit den objektseitig definierten Variablen strukturell verbunden sind, z.B. die Interpretation eines Lichtreizes (objektseitig) als "Blitz" (subjektseitig) oder die Deutung von beobachtetem Verhalten (laufen) als zielgerichtetes Handeln, "um sich vor dem Gewitter in Sicherheit zu bringen" (subjektseitig).

Subjektseitig und theoriesprachlich

Es gibt auch beobachtungssprachlich dargestellte Sachverhalte, die subjektseitig definiert sind, wie z.B. je eigene Gefühle oder Wahrnehmungen (etwa des Unterschiedes zweier Stimuli beim Weber'schen Gesetz).

Subjektseitig und beobachtungssprachlich

Schließlich gibt es auch theoriesprachlich beschriebene Sachverhalte, die objektseitig definiert sind, z.B. funktional gedeutete Sti-

Objektseitig und theoriesprachlich

muli (der ‚Lockruf‘ eines Vogels) oder funktional gedeutetes Verhalten (‚Balzgebärden‘).

Funktionales Verhalten

Ähnlich wie intentionales Handeln ist auch funktional gedeutetes Verhalten nicht direkt beobachtbares, sondern sinnhaft gedeutetes Verhalten.

Beim funktionalen Verhalten wird die Sinnhaftigkeit des Verhaltens durch Aufweis seiner Funktionalität für die Lebens- und Arterhaltung (Darwin) gleichsam aus einer Außenperspektive und aufgrund allgemeiner empirischer Gesetzmäßigkeiten des Verhaltens zum beobachteten Verhalten hinzukonstruiert (z.B. in der Biologie, Tierpsychologie).

Intentionales Handeln

Beim intentionalen Handeln wird die Sinnhaftigkeit des Verhaltens dagegen durch seine Verbindung mit individuellen Sinngehalten gleichsam aus der Innenperspektive der Vpn hinzukonstruiert (z.B. in der Sozialpsychologie, Psychotherapie).

3.1.1 Exkurs: Materielle und strukturelle Reduktion

Damit besteht bereits ein grundlegender Unterschied zwischen Physik und (Verhaltens-)Biologie. Als Erklärungsmodell dient in der Biologie zwar ebenso wie in der Physik das deduktiv-nomologische (oder das induktiv-statistische) Erklärungsschema. Durch die Berücksichtigung der Funktionalität ändert sich aber etwas in der Art und Weise, *wie* die Biologie zu ihren Gesetzeshypothesen kommt.

Semantik der Stimuli

Das Konstruktionsprinzip für Gesetzesaussagen ist in der Biologie – anders als in der Physik – das Prinzip der strukturellen Reduktion (Bischof, 1981), bei dem nicht Materialqualitäten (d.h. messbare Eigenschaften der Stimuli) am Anfang des Konstruktionsprozesses stehen, sondern die Semantik der Stimuli, d.h. ihre Bedeutung für das System. Diese wird als evolutionsbiologische Funktionalität im Darwin’schen Sinne verstanden.

Materielle Reduktion

In der Physik erhält man die Gesetzesaussagen durch eine Form der Abstraktion, die Bischof als materielle Reduktion bezeichnet. Die allgemeinen Aussagen werden gewonnen,

- indem nur die materiellen Eigenschaften der Stimuli Beachtung finden (z.B. Reizintensität),
- während man von den strukturellen Besonderheiten der Sti-

muli (z.B. Sinnesmodalität) absieht (= abstrahiert).

Diese strukturellen Besonderheiten müssen als Randbedingungen der Gesetzesaussagen wieder eingeführt werden, wenn man die spezielleren aus den allgemeinen Gesetzen zurückgewinnen will. So ist z.B. auch im Weber'schen Gesetz

$$\Delta S = c \times S$$

die Konstante c für verschiedene Sinnesmodalitäten verschieden.

Schon in der Biologie sind der materiellen Reduktion jedoch enge Grenzen gesetzt, wofür Bischof (1981) die Motivationstheorie von Schneirla (1959) als Beispiel anführt. Diese Theorie besagt,

Grenzen materieller Reduktion in Biologie und Psychologie

- dass jeder Art motivierten Verhaltens eine Polarität von Annäherung vs. Rückzug zugrunde liegt,
- wobei Annäherung durch schwache, Rückzug durch starke Reizintensität ausgelöst wird.

Das erscheint auf den ersten Blick plausibel, wenn wir etwa daran denken, wie wir in ein zu leises Radio (schwache Reizintensität) förmlich hineinkriechen (Annäherung), um besser hören zu können, und wie wir zurückschrecken (Rückzug), wenn wir das Radio einschalten und der Lautstärkeregel zu hoch gedreht ist (starke Reizintensität).

Dennoch konnte Schneirlas Motivationstheorie einer empirischen Prüfung nicht standhalten. So wird z.B. der friedliche Schläfer durch Donnerrollen weniger leicht aufgeschreckt als durch das viel leisere Knacken neben seinem Kopf. Die aversive Reaktion ist zumindest in diesem Fall nicht einfach von der Stärke des Reizes abhängig, und Bischof (1981, 32) schließt daraus, dass die materielle Reduktion versagt hat.

Als Ausweg schlägt Bischof die – für die Biologie charakteristische – Methode der strukturellen Reduktion vor und argumentiert wie folgt:

Strukturelle Reduktion

- Beunruhigend (aversiv) ist nicht ein Reiz als solcher, sondern dessen Quelle, und zwar dann, wenn es sich bei dieser um irgendein Wesen handelt, das in der Lage wäre, mir zu schaden.
- Hierzu ist es erforderlich, dass es sich in der Nähe befindet. Daraus ergibt sich die Frage: Kann man Nähe hören?
- Die absolute Schallenergie jedenfalls ist kein sicherer Indikator für Nähe, denn der Unhold kann sich ja bemühen, leise aufzu-

treten.

- Es gibt aber einen physikalischen Zusammenhang, den er nicht beeinflussen kann: Höhere Frequenzen werden durch das schallübertragende Medium stärker gedämpft als tiefe.
- Also enthält die spektrale Zusammensetzung des Schalls Informationen über die Nähe der Quelle.
- Je größer der relative Anteil hoher Frequenzen ist, desto näher ist die Schallquelle und desto alarmierender sollte daher der Schallreiz sein.

Dieser Zusammenhang lässt sich tatsächlich auch empirisch bestätigen.

Zusammenfassung

Als Ordnungsprinzip dienen der (Verhaltens-)Biologie nicht die physikalischen (materiellen) Eigenschaften der Stimuli, sondern ihre Semantik, d.h. die Bedeutung der Reize für den Organismus.

Diese Bedeutung wird im Darwin'schen Sinne verstanden als Funktionalität für die Lebens- und Arterhaltung, nicht jedoch als Intentionalität im Sinne individueller Ziel- oder Zwecksetzungen.

3.1.2 Subjektseitig definierte Beschreibungsprädikato- ren

Grenzen der strukturellen Reduktion in der Psychologie

Aufgrund seiner hoch ausgebildeten Symbolsprache ist menschliches Verhalten weit variabler als das der Tiere.

- Menschen können sich auch individuelle Ziele setzen, die im Darwin'schen Sinne disfunktional sind (z.B. Bewußtseinserweiterung durch Drogenmißbrauch);
- Menschen haben ein breites Instrumentarium an technischen Fähigkeiten, die sie absichtsvoll zur Zielerreichung einsetzen können; und
- Menschen setzen sich ihre Ziele und wählen ihre Mittel aufgrund von komplexen Beurteilungsleistungen.

In weiten Gebieten der Psychologie ist daher auch der Zusammenhang zwischen objektseitig beschriebenen Umweltgegebenheiten und objektseitig beschriebenen Verhaltensweisen viel loser als in der Biologie.

Um dem gerecht zu werden, bedürfen wir in der Psychologie einer Vielzahl an subjektseitig definierten Prädikatoren, wobei sich unterscheiden lässt zwischen

- (subjektseitig definierten) Beschreibungstermini (z.B. Wahrnehmungen, Emotionen etc.),
- Deutungstermini (z.B. Ziel, Mittel) und
- Reflexionstermini (z.B. Werte, Einstellungen, Interessen).

Arten von subjektseitig definierten Sachverhalten

Ebenso wie die Gemeinsamkeit der Verwendung von Wörtern zur Beschreibung beobachtbarer Gegenstände wird auch die gemeinsame Verwendung von Wörtern zur Beschreibung des subjektiven Erlebens im Rahmen der Einsozialisierung in die Sprachgemeinschaft sichergestellt (vgl. Hartmann & Werbik, 2001). Während wir die objektseitig definierten Beobachtungsprädikatoren in elementaren Handlungszusammenhängen anhand von Beispielen und Gegenbeispielen einüben und ihren Gebrauch dann durch Prädikatorenregeln absichern konnten, erlernen wir die subjektseitig definierten Beschreibungsprädikatoren jedoch in relativ komplexen Handlungszusammenhängen.

Subjektseitig definierte Beschreibungsprädikatoren

Die Rede von der (subjektiven) Wahrnehmung von Sachverhalten in der uns umgebenden (objektiven) Wirklichkeit wird erst möglich, wenn die Prädikatoren, mittels derer dieser Sachverhalt dargestellt wird, bereits eingeübt sind. Z.B. kann ich erst dann sagen, "*Ich sehe* eine Blume", wenn mir das Wort ‚Blume‘ bereits zur Verfügung steht.

Wahrnehmungen

Während die objektseitig definierten Beobachtungsprädikatoren (z.B. Blume) jedoch in einer Konsenssituation eingeübt wurden, wird die subjektivistische Relativierung von Aussagen als je eigene Wahrnehmung der Wirklichkeit in Dissenssituationen erlernt,¹ z.B. in einem Dialog der folgenden Art:

Mutter: "Dort ist eine Blume."

Kind: "Nein, da ist keine Blume."

Mutter: "Doch, da ist eine Blume, aber du siehst sie nicht. Ich sehe die Blume. Ich zeige sie dir." (Zeigt auf die Blume)

Kind: "Ja, da ist eine Blume."

1. Hartmann & Werbik (2001) weisen in diesem Zusammenhang darauf hin, dass erst das unter bestimmten Umständen unvermeidliche Verfehlen einer Einigung zur Etablierung der Rede von den Wahrnehmungen führt, die eine Person (im Unterschied zu einer anderen) gemacht oder eben nicht gemacht hat. "Würden sich alle immer einig sein, wäre die Rede von Wahrnehmungen überflüssig".

Mutter: "Jetzt siehst du die Blume auch."

Oder in einem Dialog wie diesem:

Mutter: "Hier ist dein Bauklotz." (Zeigt darauf)

Kind. "Ja, da ist mein Bauklotz."

Mutter: (Stülpt eine Schachtel darüber): "Jetzt siehst du ihn nicht mehr."

Kind: "Der Bauklotz ist weg."

Mutter: "Du siehst ihn nicht. Aber der Bauklotz ist noch da."
(Nimmt die Schachtel weg): "Hier ist der Bauklotz.
Jetzt siehst du ihn." (Stülpt die Schachtel darüber):
"Und jetzt siehst du ihn nicht mehr ..."

Worauf es bei den Einführungssituationen für Wahrnehmungsprädikatoren ankommt, ist, dass die Geltung einer objektseitig definierten beobachtungssprachlichen Elementaraussage ("Hier ist ...") von ihrer unmittelbaren sinnlichen Bestätigung ("Ich sehe ...") abgekoppelt und so die Unterscheidung zwischen Aussagen über die Außenwelt ("Hier ist ...") und Aussagen über das erkennende Subjekt ("Ich sehe ...", "Du siehst ...") eingeübt werden kann.

Der Terminus ‚Wahrnehmung‘ kann dann als Oberbegriff für die Wahrnehmungsprädikatoren ‚sehen‘, ‚hören‘, ‚riechen‘, ‚schmecken‘ usw. eingeführt werden.

Emotionen

Die Prädikatoren, mit denen wir unsere Emotionen beschreiben, werden zwar ebenso wie die objektseitig definierten Beobachtungsprädikatoren in Konsenssituationen eingeübt, die Komplexität der Einführungssituation ergibt sich hier aber daraus, dass wir es nicht mit Beobachtungen der uns umgebenden Wirklichkeit zu tun haben, sondern mit nach innen gerichteten Beobachtungen, die allerdings mit objektseitig definierten Sachverhalten durch teils strukturelle, teils empirische Gesetzmäßigkeiten verbunden sind.

So bezeichnet man z.B. das Gefühl, welches wir erleben, wenn wir uns etwas sehr gewünscht und sicher damit gerechnet haben, und es dann doch nicht eingetroffen ist, als Enttäuschung.

Das ist aber *keine Definition* des Prädikators ‚Enttäuschung‘, sondern lediglich eine Prädikatenregel, welche

- den subjektseitig definierten Sachverhalt "Ich bin enttäuscht"
- mit dem ebenfalls subjektseitig definierten Sachverhalt "Ich habe mir das sehr gewünscht und sicher damit gerechnet"
- und dem objektseitig definierten Sachverhalt "Dann ist es

doch nicht eingetroffen"

strukturell verbindet.

Was das Gefühl der Enttäuschung ist, wissen wir erst, wenn wir es erlebt haben. Um also den Prädikator ‚Enttäuschung‘ zu erlernen, müssen wir lernen, ihn mit einem bestimmten inneren Zustand zu verbinden.

Dies wird möglich, weil dieser innere Zustand einen von außen beobachtbaren Ausdruck hat, d.h. durch eine empirische Gesetzmäßigkeit mit bestimmten, objektseitig definierten Reaktionen verbunden ist, z.B. mit dem Hängenlassen der Mundwinkel.

Die Situation, in der ein Kind den Prädikator ‚Enttäuschung‘ erlernen kann, sieht also z.B. folgendermaßen aus:

1. Das Kind wünscht sich etwas und rechnet sicher damit.
2. Die Mutter weiß das.
3. Der Wunsch geht nicht in Erfüllung.
4. Das Kind lässt die Mundwinkel hängen.
5. Die Mutter sagt: "Du bist jetzt enttäuscht."

Schon die Einführung von Prädikatoren, mit denen wir Emotionen beschreiben, setzt damit eine intentionale Deutung des beobachteten Verhaltens sowie ein gewisses Erfahrungswissen voraus (z.B. dass das Hängenlassen der Mundwinkel oft Ausdruck von Enttäuschung ist).

Die Perspektivität der subjektseitig definierten Beschreibungsprädikatoren

Letzteres gilt auch für die Einführung der Wahrnehmungstermini, die ohne ein Wissen darüber, ob bzw. unter welchen Bedingungen ein Gegenstand (die Blume, der Bauklotz) für andere (das Kind) wahrnehmbar ist, nicht gelingen kann.

Entsprechend dieser Theoriehaltigkeit bereits der Einführungssituation können die subjektseitig definierten Beschreibungsprädikatoren nicht einfach der Beobachtungs- oder der Theoriesprache zugerechnet werden, sondern sie werden je nach Perspektive beobachtungssprachlich oder theoriesprachlich verwendet. In Aussagen über die je eigenen Wahrnehmungen und Gefühle ("Ich sehe"; "Ich bin enttäuscht") werden sie beobachtungssprachlich verwendet. In Aussagen über die Wahrnehmungen oder Gefühle eines anderen ("Du siehst"; "Du bist enttäuscht") werden die subjektseitig definierten Beschreibungsprädikatoren dagegen theoriesprachlich verwendet.

3.1.3 Verhalten, Handlung und Widerfahrnis

Mit der Verwendung intentionaler Erklärungen entfernen wir uns in der Psychologie noch einen Schritt weiter von den ‚harten‘ Naturwissenschaften. Geht schon die Verhaltensbiologie über das Beobachtbare hinaus, indem sie Verhalten *funktional* deutet, so gehen wir noch einen Schritt weiter und deuten beobachtbares Verhalten als *intentionale Handlung*.

Dies wirft eine Reihe von Überprüfungsproblemen auf, wie z.B. die Frage, ob wir uns damit überhaupt noch auf dem Boden einer empirischen Wissenschaft bewegen bzw. ob und ggf. wie intentionale Erklärungen an der Erfahrung scheitern können.

Um diese Fragen näher untersuchen zu können, müssen wir zunächst unsere Terminologie um Handlungs- und Deutungstermini erweitern.

Wie wir bereits festgestellt haben, gehören Handlungstermini ebenfalls zu den Beschreibungstermini, d.h. dass sie (in einem ersten Schritt) anhand von Beispielen und Gegenbeispielen eingeübt werden können. Beispiele für Handlungen sind dabei gleichzeitig Gegenbeispiele für Widerfahrnisse (vgl. Tab. 3.1.2).

Handlung und
Widerfahrnis

Tab. 3.1.2: Beispiele für Handlungen und Widerfahrnisse

Handlung	Widerfahrnis
Klavierspielen	Niesen
Sich zu Boden werfen	Hinfallen
Demonstrieren	Erwachen
(etwas) ignorieren	(etwas) übersehen
Reden	Beleidigt werden
Jemanden schlagen	Geschlagen werden

Die Termini ‚Handlung‘ und ‚Widerfahrnis‘ stehen also in einem konträren Gegensatz zu einander:

$$x \in \text{Handlung} \Rightarrow x \notin \text{Widerfahrnis}$$

$$x \in \text{Widerfahrnis} \Rightarrow x \notin \text{Handlung}$$

Handlung und
Verhalten

Beispiele für Handlungen sind zugleich Beispiele für Verhalten:

$$x \in \text{Handlung} \Rightarrow x \in \text{Verhalten},$$

aber nicht alle Verhaltensweisen sind zugleich Handlungen. Auch manche Beispiele für Widerfahrnisse sind Beispiele für Verhalten.

Dies erlaubt uns, Handlungen von bloßem Verhalten zu unterscheiden:

$$x \in \text{blo\ss es Verhalten} \Leftrightarrow x \in \text{Verhalten} \wedge x \in \text{Widerfahrnis}$$

	Verhalten
Widerfahrnis (blo\ss es Verhalten)	Niesen Hinfallen Erwachen (etwas) \u00fcbersehen
Handlung	Klavierspielen Sich zu Boden werfen Demonstrieren (etwas) ignorieren Reden Jemanden schlagen

Tab. 3.1.3: Beispiele f\u00fcr Handlungen und blo\ss es Verhalten

Verhalten ist also ein Oberbegriff f\u00fcr Handlung und blo\ss es Verhalten (vgl. Tab. 3.1.3). Dabei ist blo\ss es Verhalten ein Verhalten, das mir lediglich widerf\u00e4hrt. Eine Handlung ist dagegen *beabsichtigt*. Eine Handlung ist ein Verhalten, das mir nicht nur widerfahren ist, sondern zu dem ich mich quasi *aufgefordert* habe.

3.1.4 Ziel und Mittel

Ausgehend vom Terminus ‚Aufforderung‘ k\u00f6nnen wir jetzt die grundlegenden Deutungstermini definieren, welche f\u00fcr intentionale Erkl\u00e4rungen ben\u00f6tigt werden. Zun\u00e4chst unterscheiden wir zwischen afinalen und finalen Aufforderungen:

Afinale und finale Aufforderungen

- *Afinale* Aufforderungen sind Aufforderungen zur Ausf\u00fchrung einer Handlung, z.B. "\u00d6ffne das Fenster!"
- *Finale* Aufforderungen sind Aufforderungen zur Herbeif\u00fchrung eines Sachverhalts, z.B. "Lass frische Luft in den Raum!"

Hat man gelernt, dass bestimmte Handlungen (unter entsprechenden Bedingungen) regelm\u00e4\u00dfig das Eintreten bestimmter Sachverhalte nach sich ziehen, so kann man auch lernen, afinale Aufforderungen mit finalen Aufforderungen zu verbinden: z.B. " \u00d6ffne das Fenster, *um* frische Luft in den Raum *zu* lassen!" Wir k\u00f6nnen dann auch sagen:

Mittel und Zweck

- Die Handlung H hat den *Zweck* Z.
- Die Handlung H ist ein *Mittel* zur Herbeif\u00fchrung von Z.

Ist der bezweckte Sachverhalt eingetroffen, so ist die Handlung abgeschlossen. Der Zweck ist damit aufgehoben, der Zweck entfällt. Die Handlung hat ihr Ziel erreicht.¹

Kontrollfragen

- Geben Sie Beispiele für
 - a) objektseitig definierte Sachverhalte, die beobachtungssprachlich beschrieben werden,
 - b) subjektseitig definierte Sachverhalte, die beobachtungssprachlich beschrieben werden,
 - c) objektseitig definierte Sachverhalte, die theoriesprachlich beschrieben werden,
 - d) subjektseitig definierte Sachverhalte, die theoriesprachlich beschrieben werden!
- Erläutern Sie die Gemeinsamkeit und den Unterschied zwischen funktionalem Verhalten und intentionalem Handeln!
- Erläutern Sie den Unterschied zwischen materieller und struktureller Reduktion!
- Erläutern Sie, was Bischof in diesem Zusammenhang unter
 - a) den Materialqualitäten und
 - b) der Semantik der Stimuli versteht!
- Erläutern Sie am Beispiel des Weber'schen Gesetzes das Prinzip der materiellen Reduktion!
- Stellen Sie die Motivationstheorie von Schneirla dar und geben Sie Beispiele und Gegenbeispiele für ihre Geltung!
- Erläutern Sie im Kontrast zur Motivationstheorie von Schneirla das Prinzip der strukturellen Reduktion!
- Welche Grenzen sind der strukturellen Reduktion in der Psychologie gesetzt?
- Warum ist der Zusammenhang zwischen objektseitig definierten Umweltgegebenheiten und objektseitig beschriebenen Verhaltensweisen in der Psychologie weit loser als in der Biologie?
- Welche Arten von subjektseitig definierten Prädikatoren können in der Psychologie unterschieden werden? Geben Sie Beispiele!
- Wodurch unterscheidet sich die exemplarische Einführung subjektseitig definierter Beschreibungsprädikatoren von der exemplarischen Einführung objektseitig definierter Beobachtungsprädikatoren?
- Wie kann der Terminus ‚Wahrnehmung‘ eingeführt werden? Woraus ergibt sich die Komplexität der Einführungssituation?
- Erläutern Sie anhand des Terminus ‚Enttäuschung‘ die Einführung von Prädikatoren, mit denen wir unsere Emotionen beschreiben!

1. Auch in der Verhaltensbiologie können wir von zielgerichtetem Verhalten sprechen. Die Zielgerichtetheit wird dort aber aus der Aussenperspektive des Forschers zum beobachtbaren Verhalten hinzukonstruiert, indem man aufzeigt, dass das Verhalten a) eine bestimmte Funktion erfüllt, und b) regelmäßig wieder eingestellt wird, sobald es diese Funktion erfüllt hat.

- Woraus ergibt sich die Komplexität der Einführungssituation in diesem Fall?
- Worin besteht die Theoriehaltigkeit der Einführungssituation subjektseitig definierter Beschreibungsprädikatore?
- Erläutern Sie, ob bzw. inwiefern die subjektseitig definierten Beschreibungsprädikatore der Beobachtungs- oder der Theoriesprache zuzurechnen sind!
- Erläutern Sie die terminologische Unterscheidung zwischen
 - a) Handlung, Verhalten, Widerfahrnis und
 - b) Verhalten, Handlung, bloßem Verhalten!
- Erläutern Sie den Unterschied zwischen finalen und afinalen Aufforderungen!
- Erläutern Sie die Einführung der Termini ‚Mittel‘ und ‚Zweck‘!
- Was versteht man darunter, dass eine Handlung mit Eintreten des bezweckten Sachverhaltes abgeschlossen ist?

3.2

Das intentionale Erklärungsmodell

Das intentionale Erklärungsmodell

Beobachtungssprachliche Prädikatore werden in Handlungszusammenhängen erlernt. Deutungssprachliche Prädikatore (Zweck, Mittel) werden in Redezusammenhängen erlernt. Bei den Handlungsprädikatore ist beides miteinander verknüpft.

Zwischen Deutungsprädikatore und Beobachtungsprädikatore besteht ein grundlegender Unterschied. Wir beobachten nicht einfach, welches Ziel wir mit einer Handlung verfolgen oder welches Verhalten wir für ein geeignetes Mittel zur Erreichung dieses Zieles halten, sondern: Durch die Behauptung, dass ich (oder jemand anders) mit der Handlung H das Ziel Z verfolge, wird der behauptete Zusammenhang zwischen dem beobachtbaren Verhalten (V) und dem Ziel (Z), zu dessen Erreichung es als Mittel dient, erst *konstruiert*. Dies geschieht nach dem Schema des intentionalen Erklärungsmodells (v. Wright, 1974):

N beabsichtigt, Z herbeizuführen.

N meint, dass H (in der gegebenen Konstellation) *das* geeignete Mittel ist.

Also macht sich N daran, H auszuführen.

Deutungsprädikatore beschreiben somit keine Beobachtungen, sondern *theoretische Konstrukte*. Dies bewirkt spezielle Überprüfungsprobleme, die zu klären schon allein deswegen unerlässlich ist, weil das intentionale Erklärungsmodell nicht nur in der psychologischen Forschung, sondern auch in der psychologischen Praxis weite Anwendung findet, z.B. in der Psychotherapie, in der psychologischen Beratung und/oder in der Konfliktvermittlung (Mediation).

Spezialfall
Psychoanalyse

Eine gewisse Sonderrolle spielt dabei die Psychoanalyse insofern, als sie Sinngehalte zum Handeln der Klienten hinzukonstruiert (sog. unbewusste Motive), die den Klienten selbst nicht bewusst sind oder zumindest bisher nicht bewusst waren.

Den Terminus ‚Sinngelhalte‘ verwenden wir als Sammelbegriff für all jene Dinge, die Menschen in ihren Köpfen haben und die man von außen nicht beobachten kann: Absichten, Situationseinschätzungen, Meinungen und Wissen, Lösungsstrategien usw.

Sinngelhalte sind stets etwas, das zum beobachtbaren Verhalten hinzukonstruiert ist, sei es durch den Akteur selbst oder durch sein Gegenüber. Ohne das Hinzukonstruieren von Sinngelhalten auf Seiten unseres Gegenübers wäre interpersonales Verstehen gar nicht möglich.

Bei den unbewussten Motiven ist es nun so, dass der Analytiker Absichten des Klienten hinzukonstruiert, die dieser selbst *nicht* mit seinem Handeln verbindet. Zugleich erhebt der Psychoanalytiker einen Geltungsanspruch für seine Deutungen, und zwar gegen die Selbstdeutung des Klienten.

Wie dieser Geltungsanspruch eingelöst werden kann, ist das methodologische Hauptproblem der Psychoanalyse, das sich insbesondere dann ergibt, wenn sich der Klient den Interpretationen des Analytikers nicht anschließt, sondern auf seiner Selbstdeutung beharrt, was man dann als Widerstand bezeichnet.

3.2.1 Bedingungs- vs. Begründungszusammenhänge

Mit der Einführung des intentionalen Erklärungsmodells erfährt sowohl das Gegenstandsverständnis als auch das Wissensideal der Psychologie eine wesentliche Erweiterung. Im Unterschied zur deduktiv-nomologischen und zur induktiv-statistischen Erklärung geht es jetzt nicht mehr um Ursache-Wirkungs-Relationen (Bedingungs- vs. Begründungszusammenhänge), sondern um Ziel-Mittel-Relationen (Begründungszusammenhänge).

Zum Gegenstandsverständnis der Psychologie gehören jetzt z.B. nicht nur Fragen nach den Bedingungen, unter denen wir mit der gewaltförmigen Eskalation von Konflikten zu rechnen haben, sondern auch Fragen danach, *warum* bestimmte Konflikte gewaltförmig eskalieren. Ein Beispiel

Eine angemessene Antwort darauf könnte u.a. lauten, dass die Konzeption von Konflikten als win-lose-Situationen, gepaart mit mangelnder Nachgiebigkeit und dem Fehlen gewaltfreier Mittel zur Durchsetzung eigener Ziele, eine gewaltförmige Konflikteskalation nach sich ziehen.

Die Konzeption von Konflikten als win-lose-Situationen (A), die mangelnde Nachgiebigkeit der Konfliktparteien (B) und das Fehlen gewaltfreier Mittel zur Durchsetzung je eigener Ziele (C) stellen in dieser Erklärung jedoch nicht einfach Randbedingungen (unabhängige Variablen) dar, aus welchen – wie bei der deduktiv-nomologischen Erklärung – qua universeller empirischer Subjunktion

$$\forall x. \{A(x) \wedge B(x) \wedge C(x)\} \rightarrow D(x)$$

auf die Konflikteskalation (D) als abhängige Variable geschlossen wird. Die Erklärung ist nicht empirischer, sondern struktureller Art und bedient sich des intentionalen Erklärungsmodells zur Konstruktion einer sachlogischen Relation zwischen den Variablen.

Wenn jemandem eine bestimmte Absicht und eine bestimmte Mittelmeinung unterstellt wird, und er verhält sich nicht entsprechend, dann wird dadurch nicht die Zusammenhangsbehauptung falsifiziert, dass aus einer bestimmten Absicht bei bestimmtem Mittelwissen eine bestimmte Handlung resultiert, sondern die in der Handlungserklärung enthaltenen Prämissen: Wenn sich jemand nicht entsprechend verhält, dann hat er entweder nicht die unterstellte Absicht verfolgt, oder er hat nicht über das unterstellte Mittelwissen verfügt, oder er ist gehindert worden, oder ...

Im konkreten Fall der Konflikteskalation kann die Erklärung wie folgt konstruiert werden:

N_1 beabsichtigt, das Ziel Z herbeizuführen, und ist nicht bereit, darauf zu verzichten (mangelnde Nachgiebigkeit).

N_1 meint, sich über sein Ziel in Konflikt mit N_2 zu befinden, und dass er in diesem Konflikt nur entweder gewinnen (sich durchsetzen) oder verlieren (nachgeben und auf Z verzichten) kann (win-lose).

Also macht sich N_1 daran, sein Ziel Z gegen N_2 durchzusetzen (Aggression).

Und weiter:

N_1 beabsichtigt, sein Ziel Z gegen N_2 durchzusetzen.

Unter den Mitteln, die N_1 dafür zur Verfügung stehen, scheinen ihm nur die gewaltförmigen Mittel erfolversprechend (z.B. weil N_2 nicht verhandlungsbereit ist).

Also macht sich N_1 daran, Gewalt anzuwenden.

Obwohl wir es dabei nicht mit einer empirischen, sondern mit einer sachlogischen Erklärung zu tun haben, wird empirische Forschung dadurch nicht überflüssig. Im Gegenteil wirft die sachlogische Erklärung der Konflikteskalation eine ganze Reihe von empirischen Anschlussfragen auf, zum Beispiel, unter welchen Bedingungen Konflikte als win-lose-Situationen interpretiert werden, welche Bedingungen Unnachgiebigkeit erzeugen oder welche Bedingungen gewaltfreie Mittel als aussichtslos erscheinen lassen etc.

3.2.2 Die Hypothesenförmigkeit intentionaler Erklärungen

Zwischen dem intentionalen und dem deduktiv-nomologischen Erklärungsmodell bestehen einige grundlegende Unterschiede:

Unterschiede zwischen dem intentionalen und dem deduktiv-nomologischen Erklärungsmodell

- Die deduktiv-nomologische Erklärung beruht auf der Behauptung einer universellen *empirischen* Gesetzmäßigkeit, die stets *hypothetisch* bleibt (selbst wenn sie sich noch so gut bewährt hat).
- Die intentionale Erklärung beruht dagegen nicht auf einer empirischen Gesetzmäßigkeit, sondern auf einem *sachlogischen* Zusammenhang, der nicht bloß hypothetisch ist, sondern *mit Sicherheit* behauptet werden kann.

Dieser Zusammenhang folgt daraus, wie die Mittel-Ziel-Relation konstruiert ist. Denn die Deutungstermini (Ziel, Mittel) beschreiben keine Selbstbeobachtungen, sondern theoretische Konstrukte, die zum beobachtbaren Verhalten (auch zum eigenen) nach bestimmten Regeln hinzukonstruiert werden.

- Aufgrund dieser Regeln kann ich z.B. nicht behaupten, dass ich mit einem bestimmten Verhalten ein bestimmtes Ziel verfolge und gleichzeitig behaupten, dass ich dieses Ziel gar nicht erreichen will; und
- ich kann auch nicht behaupten, dass ich mit einem bestimmten Verhalten ein bestimmtes Ziel verfolge und gleichzeitig behaupten, dass ich dieses Verhalten für kein geeignetes Mittel halte, um dieses Ziel zu erreichen.

Wenn ich solches dennoch behauptete, wird damit nicht der sachlogische Zusammenhang widerlegt, der im intentionalen Erklä-

rungsmodell angesprochen ist, sondern lediglich meine Sprachkompetenz.

- Das deduktiv-nomologische Erklärungsmodell erklärt konkretes Verhalten unter Verweis auf objektseitig gegebene Randbedingungen, die ebenfalls konkret beobachtbar sind und deren Bestehen oder Nichtbestehen *cum grano salis* mittels Beobachtung entschieden werden kann.
- Das intentionale Erklärungsmodell erklärt konkretes Verhalten dagegen unter Verweis auf subjektseitig gegebene Randbedingungen (Absichten, Mitteleinschätzung), die nicht direkt beobachtbar sind und deren Bestehen zwar anhand von Beobachtungen widerlegt, aber nicht definitiv bewiesen werden kann, die also stets *hypothetisch* bleiben.

Diese Hypothesenförmigkeit von Aussagen über Sinngehalte folgt daraus, dass Sinngehalte nicht unmittelbar beobachtbar sind, sondern aus dem beobachtbaren Verhalten mittels des intentionalen Erklärungsmodells erschlossen werden, wobei wir es quasi mit einer Gleichung mit zwei Unbekannten (Z und V) zu tun haben:

N beabsichtigt Z, herbeizuführen.

N meint, dass V (in der gegebenen Konstellation) *das* geeignete Mittel ist.

Also macht sich N daran, V auszuführen.

Das Überprüfungsproblem verschärft sich noch dadurch, dass Handlungen nicht nur die intendierten Wirkungen haben, sondern oft auch unbeabsichtigte Nebenfolgen. Hinzu kommt, dass Menschen nicht aufgrund der objektiven Beschaffenheit der Dinge in ihrer Umwelt handeln, sondern aufgrund der Bedeutung, die diese für sie haben (Blumer, 1973).

Handlungsabläufe

Welche Konsequenzen dies hat, wird deutlich, sobald wir nicht nur isolierte Handlungen, sondern Handlungsabläufe betrachten (vgl. Abb. 3.2.1).

Die objektseitig definierten Umweltgegebenheiten konstituieren sich zur subjektseitig definierten Situation, indem sie von den Subjekten im Lichte ihrer je eigenen Biographie, ihrer aktuellen Bedürfnisse, ihrer vorangegangenen Handlungen usw. interpretiert werden. Den Ausschnitt aus der Situation, der im konkreten Fall handlungsrelevant wird, bezeichnen wir im Anschluss an Hoyningen-Huene (1983) als handlungsauslösende Konstellation.

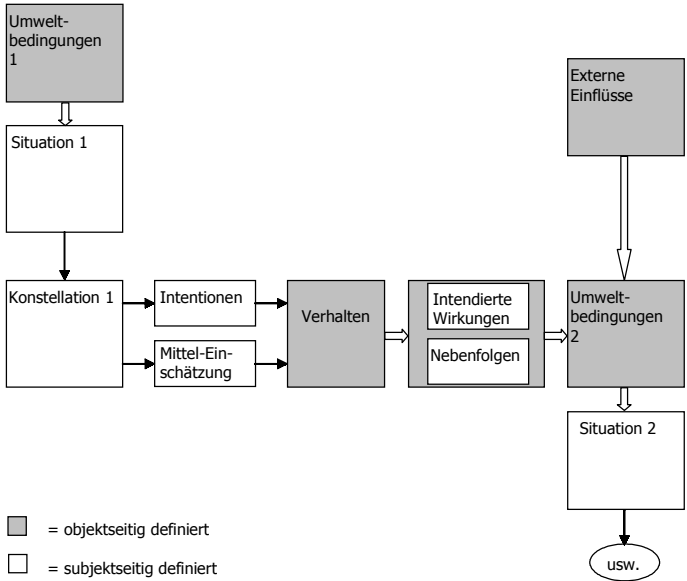


Abb. 3.2.1: Schematische Darstellung von Handlungsabläufen

Die daraus folgenden Intentionen und Mitteleinschätzungen sind ebenfalls subjektseitig definiert.

Das Verhalten der Subjekte, dessen Wirkungen und Folgen, sowie die – im Wechselspiel mit externen Einflüssen – daraus resultierende Veränderung der Umweltgegebenheiten sind objektseitig definiert, wogegen die Unterscheidung der Wirkungen und Folgen in intendierte Wirkungen und nicht-intendierte Nebenfolgen subjektseitig definiert ist.

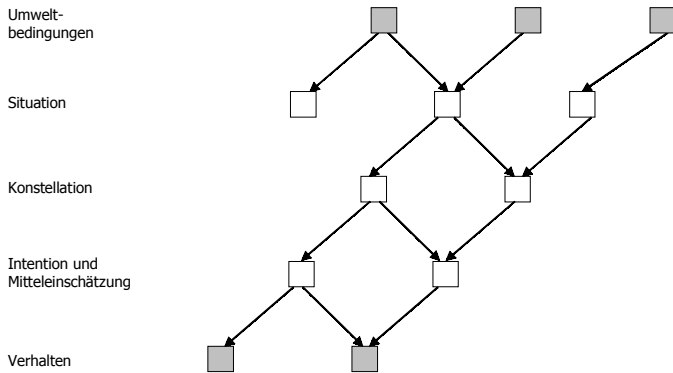
Daraus ergibt sich eine grundlegende Interpretationsoffenheit von Handlungen und Sinngehalten, wie sie in Abb. 3.2.2 dargestellt ist, und eine intentionale Erklärung wird daher (außer in artifiziellen Experimenten) nie zwingend sein.

Die Interpretationsoffenheit von Handlungen und Sinngehalten

Dieselben Umweltbedingungen können von den Vpn als verschiedene Situationen interpretiert werden, und unterschiedliche Umweltbedingungen können als dieselbe(n) (bzw. als vergleichbare) Situation(en) interpretiert werden.

Verschiedene Ausschnitte derselben Situation können zur handlungsauslösenden Konstellation werden, und dieselbe Konstellation kann sich aus verschiedenen Situationen ergeben.

Abb. 3.2.2: Die Interpretationsoffenheit von Handlungen und Sinngehalten



In derselben Konstellation können verschiedene Intentionen und/oder Mittelmeinungen zum Tragen kommen, und dieselben Intentionen und Mittelmeinungen können in verschiedenen Konstellationen zum Tragen kommen.

Aus denselben Intentionen und Mittelmeinungen kann verschiedenes Verhalten resultieren, und dasselbe Verhalten kann sich aus verschiedenen Intentionen und Mittelmeinungen ergeben.

3.2.3 Die empirische Prüfbarkeit intentionaler Erklärungen

Das Kriterium der Falsifizierbarkeit

Die Interpretationsoffenheit von Handlungen und Sinngehalten lässt sofort Zweifel daran entstehen, ob das intentionale Erklärungsmodell mit dem Selbstverständnis der Psychologie als einer Erfahrungswissenschaft verträglich ist bzw. genauer: ob Hypothesen über subjektseitig definierte Handlungsprämissen überhaupt an der Erfahrung scheitern können.

Der Behaviorismus hat dies rundweg verneint. Für den Behaviorismus galt nur das einer wissenschaftlichen Behandlung zugänglich, was man direkt beobachten kann. Das sind einerseits die objektseitig definierten Umweltbedingungen (Stimuli) und andererseits das objektseitig definierte Verhalten (Response). Alles andere wurde als "private events" (Buss, 1961) denunziert und aus der Psychologie verbannt.

Seit der kognitiven Wende gilt diese Meinung zwar als überholt, und es hat sich die Auffassung durchgesetzt, dass man dem Ge-

genstand der Psychologie nicht gerecht werden kann, wenn man die Sinngehalte aus der Psychologie ausklammert. Es bleibt aber doch erst noch zu zeigen, dass die Psychologie ihre Erfahrungswissenschaftlichkeit damit nicht etwa aufgegeben hat, sondern dass intentionale Erklärungen trotz der Interpretationsoffenheit von Handlungen und Sinngehalten empirisch falsifizierbar sind.

In diesem Kontext kommt der Frage, welche Aspekte der Situation zur handlungsauslösenden Konstellation gehören, eine Schlüsselrolle zu. Diese ist nämlich nicht als Frage nach einem empirischen Prozess aufzufassen, in dem sich bestimmte Situationsaspekte gleichsam zur Konstellation verdichten, sondern als die Frage danach, welche Aspekte der Situation im Rahmen einer intentionalen Erklärung benötigt werden und von welchen abgesehen (= abstrahiert) werden kann (vgl. Kempf, 1999).

Die handlungsauslösende Konstellation

Was die Konstellation ausmacht, sind jene Aspekte der Situation, die es einerseits erlauben, zwischen intentionalen (= beabsichtigten), kontraintentionalen (= der Intention entgegenlaufenden), paraintentionalen (= irrtümlichen) und periintentionalen (= in Kauf genommenen) Folgen einer Handlung zu unterscheiden (vgl. Brandtstädter & Greve, 1999), und die es andererseits erlauben, die Responsivität der Handlung darzustellen, d.h. erstens herauszuarbeiten, worauf die Handlung antwortet, und zweitens aufzuzeigen, dass sie eine (vom Standpunkte des Subjektes aus gesehen) *sinnvolle* Antwort ist (vgl. Waldenfels, 1999). Mit anderen Worten: Es geht darum, zu zeigen, dass die Person getan hat, was aus ihrer Sicht in dieser Situation zu tun war.

Damit ist schon eine Teilantwort auf die Frage angedeutet, welche Art von empirischer Wahrheit intentionale Erklärungen beanspruchen können bzw. worin die Wahrheitskriterien bestehen, an denen sie zu messen sind. Wenn wir nämlich die Frage stellen, welche Situationsaspekte im Rahmen einer intentionalen Erklärung benötigt werden, so können wir diese nicht beantworten, ohne vorher geklärt zu haben, *wofür* sie denn benötigt werden sollen. Ein abstraktes Streben nach Wahrheit als Selbstzweck kann hier keine tragfähige Antwort geben, wohl jedoch das Interesse an einer erfolgreichen Praxis – z.B. der Konfliktbewältigung. Die Richtigkeit einer intentionalen Erklärung erweist sich dann letztlich am Erfolg der Konfliktlösungsbemühungen, welche darauf gegründet sind.

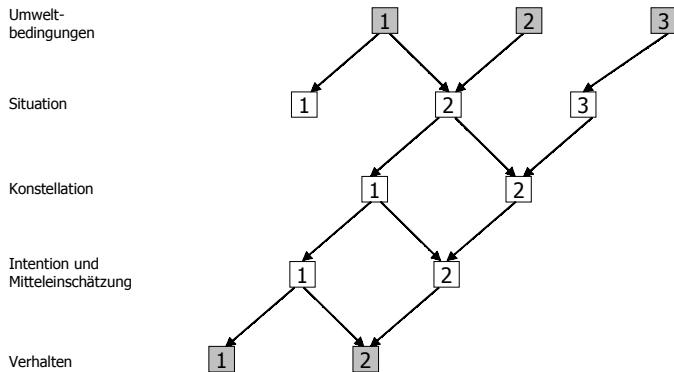
Wahrheitskriterien

Gleichwohl gilt es festzuhalten, dass intentionale Erklärungen einen nicht aufzulösenden hypothetischen Charakter besitzen. Dasselbe trifft (mangels Verifizierbarkeit empirischer Allgemeinaussagen) aber auch für naturwissenschaftliche Erklärungen im deduktiv-nomologischen Erklärungsschema zu. Entscheidend für den Empiriebezug wissenschaftlicher Erklärungen ist jedoch nicht deren hypothetischer Charakter, sondern die Frage, ob die Hypothesen an der Erfahrung scheitern können. Für intentionale Erklärungen ist dies in dreierlei Hinsicht gegeben:

Genetische Argumentation

Erstens: Hypothesen über Handlungsgründe können als falsifiziert gelten, wenn die Entstehung (Genese) der zur Erklärung des beobachteten Verhaltens unterstellten Sinngehalte nicht von oben (= von den gegebenen Umweltbedingungen her) rekonstruiert werden kann. Im Anschluss an Schwemmer (1976) sprechen wir dann vom Scheitern der genetischen Argumentation.

Abb. 3.2.3: Scheitern der genetischen Rekonstruktion



So kann in Abb. 3.2.3 das Zustandekommen des Verhaltens 1 sowohl aus der Umweltbedingung 1 als auch aus den Umweltbedingung 2 hervorgegangen sein, nicht jedoch aus der Umweltbedingung 3, weil diese nicht mit den unterstellten Handlungsgründen (Intentionen und Mitteleinschätzung) vereinbar ist, welche wir zur Erklärung des Verhaltens unterstellen müssen. Sind diese Umweltbedingungen (3) gegeben, dann liefern die angenommenen Sinngehalte daher keine brauchbare Erklärung des Verhaltens. Die Erklärung ist gescheitert.

- Z.B. streiten sich zwei Schwestern um eine Orange (Verhalten 1).
- Wir versuchen, dies dadurch zu erklären, dass jede der beiden

gerne eine Orange essen will (Handlungsgründe 1).

- Es sind aber genügend Orangen für beide da (Umweltbedingung 3).
- Also muss der Streit einen anderen Grund haben: Der Wunsch, eine Orange zu essen (allein), ist keine brauchbare Erklärung für den Streit.

Wenn wir unsere Handlungserklärung dennoch aufrecht erhalten wollten, müssten wir annehmen, dass zwar genügend Orangen da sind, dass die beiden Schwestern dies jedoch nicht wussten.

Es genügt aber nicht, einfach zu behaupten, dass sie dies nicht wussten, sondern es sind Gründe vorzubringen, *warum* sie es nicht wussten bzw. nicht wissen konnten.

Bezogen auf unser obiges Schema bedeutet dies, dass eben doch ein möglicher Pfad von der Umweltbedingung (3) zu der Situationseinschätzung (2) führt, welche wir in unserer Handlungserklärung unterstellen. Gelingt dies, so war die genetische Argumentation erfolgreich, und wir können die Handlungserklärung (vorläufig) als Hypothese beibehalten. (vgl. Abb. 3.2.4).

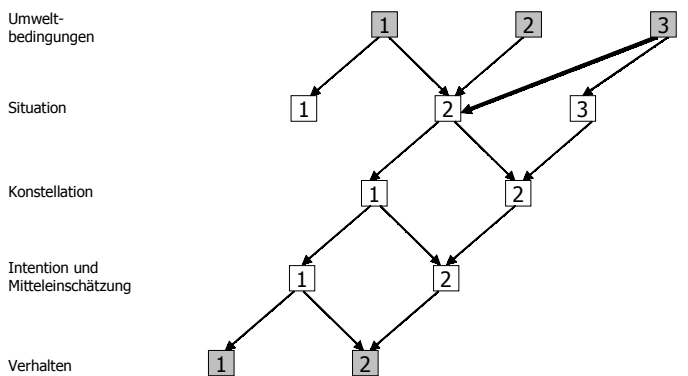
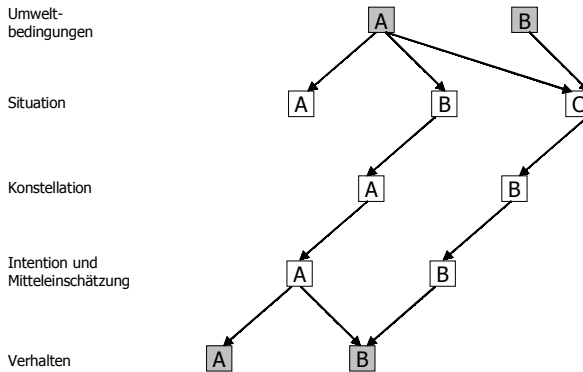


Abb. 3.2.4: Erfolgreiche genetische Argumentation

Die allgemeine Form der genetischen Argumentation ist dabei die Form einer Narration, welche die Entstehung der handlungsauslösenden Konstellation in Form einer Geschichte darstellt. In unserem Beispiel könnte sie etwa so aussehen, dass die Mutter Orangen gekauft und in den Keller getragen hat, während die beiden Schwestern noch in der Schule waren. Aus der Schule heimgekommen, stürzen sich die Schwestern sofort auf die eine Orange, die noch vom Vortag in der Küche liegt, und beginnen ihren Streit ...

Empirische Reduktion *Zweitens*: Intentionale Erklärungen können als falsifiziert gelten, wenn die empirische Reduktion der Handlungsgründe scheitert, d.h. wenn das aufgrund der unterstellten Sinngehalte prognostizierte Verhalten nicht eintritt, so dass sich die angenommenen Handlungsgründe nicht von unten (= vom objektseitig gegebenen Verhalten und seinen Wirkungen her) rekonstruieren lassen.

Abb. 3.2.5: Scheitern der empirischen Reduktion



Wenn wir in Abb. 3.2.5 z.B. unterstellen, dass Intention und Mittelwissen B besteht, dann folgt daraus, dass das Verhalten B zu beobachten sein muss. Tritt das Verhalten nicht ein, dann können wir daraus schließen, dass die unterstellten Sinngehalte (Intentionen und Mittelwissen) keine brauchbare Erklärung darstellen. Die Erklärung ist gescheitert.

- Z.B. streiten sich zwei Schwestern um eine Orange (Verhalten B).
- Wir versuchen, dies dadurch zu erklären, dass jede der beiden gerne die Orange essen will (Handlungsgründe B).
- Als eine der beiden den Streit gewonnen hat, isst sie die Orange gar nicht, sondern wirft sie in den Müll (Verhalten A).
- Also muss der Streit einen anderen Grund haben: Der Wunsch, die Orange zu essen, ist keine brauchbare Erklärung für den Streit.

Pragmatische Evaluation

Drittens: Hypothesen über Handlungsgründe können als falsifiziert gelten, wenn ihre pragmatische Evaluation scheitert, d.h. wenn eine auf die angenommenen Handlungsgründe gestützte Praxis misslingt.

In unserem Beispiel:

- Zwei Schwestern streiten um eine Orange.
- Die Mutter versucht, in dem Streit zu vermitteln.
- Sie spricht erst mit der einen. Die sagt: "Ich habe mich schon den ganzen Tag darauf gefreut, die Orange zu essen, und ich will nichts davon abgeben."
- Dann spricht die Mutter mit der anderen. Die sagt: "Ich brauche die Orange. Ich will einen Kuchen backen, und im Rezept steht, dass ich die Schale einer Orange hineintun muss."
- Die Mutter meint, der Konflikt sei gelöst: Die eine Tochter kann die Orange essen, und die andere bekommt die Orangenschalen für ihren Kuchen.
- Als sie ihren Töchtern diese Lösung vorschlägt, will sich keine der beiden darauf einlassen. Sie streiten weiter.
- Also muss der Streit einen anderen Grund haben: Dass die eine die Orange essen will und die andere die Orangenschalen braucht, ist keine brauchbare Erklärung für den Streit.

3.2.4 Exkurs: Der Unkritisierbarkeitsverdacht gegenüber der Psychoanalyse

Sinngehalte sind Deutungen. Sie sind theoretische Konstrukte, die nach den Regeln des intentionalen Erklärungsmodells zum beobachteten Verhalten hinzukonstruiert werden. Dies kann sowohl aus der Innenperspektive durch das handelnde Subjekt selbst oder aus einer Außenperspektive durch einen unbeteiligten Beobachter (z.B. den Psychotherapeuten) geschehen. Entsprechend kann man zwischen Selbstdeutungen (des Klienten) und Fremddeutungen (des Therapeuten) unterscheiden.

Selbstdeutungen und Fremddeutungen

Stimmen die Selbstdeutungen mit den Fremddeutungen nicht überein, so kann dies verschiedene Gründe haben.

1. Die Selbstauskünfte sind falsch (gelogen).
2. Die Selbstdeutungen sind falsch (Der Klient ist sich seiner wahren Motive nicht bewusst).
3. Die Fremddeutungen sind falsch (Der Therapeut irrt sich).

Grundsätzlich besteht in der interpersonalen Kommunikation ein *Primat der Selbstdeutungen* vor den Fremddeutungen. Nur ich selbst kann wissen, was ich wirklich will. Gleichwohl kann die Wahrhaftigkeit der Selbstauskünfte und/oder die Wahrheit der Selbstdeutungen widerlegt werden, wenn sie methodisch nicht korrekt konstruiert sind (Verwicklung in Widersprüche), wenn sie

sich genetisch nicht rekonstruieren lassen, wenn ihre empirische Reduktion scheitert und/oder wenn sie der pragmatischen Evaluation nicht standhalten.

Vulgärfreudianische
Immunsierungstendenzen

Bei der Rede von unbewussten Motiven tritt dagegen ein Primat der Fremddeutungen vor den Selbstdeutungen in Kraft. Dies kann so weit gehen, dass der Psychoanalytiker ein Interpretationsmonopol bezüglich der unbewussten Motive beansprucht. Werden diese Deutungen vom Klienten nicht übernommen, so leistet er Widerstand, womit sich der Verdacht ergibt, dass psychoanalytische Deutungen nicht kritisierbar sind: Übernimmt der Klient die Fremddeutungen des Analytikers, so sind die Deutungen des Psychoanalytikers wahr, und dem Klienten sind seine (bisher) unbewussten Motive damit bewusst geworden. Hält er an seinen Selbstdeutungen fest, so leistet er Widerstand, und die Deutungen des Psychoanalytikers sind ebenfalls wahr. In diesem Sinne schreibt z.B. Eysenck (1985, 139f), dass

"... Freud eine sehr schlaue Masche [hatte], mit abweichenden Absichten fertigzuwerden. Wenn der Patient seinen Deutungen zustimmte, dann erhob er den Anspruch, dass sie zweifelsfrei richtig seien. Stimmte der Patient dagegen nicht zu, so behauptete er, dass dies in einem ‚Widerstand‘ gegen die Analyse begründet läge, welcher die Deutung gerade deshalb unannehmbar erscheinen ließ, weil sie richtig sei; daher musste die fehlende Übereinstimmung gleichfalls die Richtigkeit der Theorie anzeigen. Selbstredend gibt es damit keine Möglichkeit, die Theorie zu widerlegen".

Während Eysencks Vorwürfe zwar an Freud selbst windschief vorbeigehen, kann andererseits nicht ausgeschlossen werden, dass manche Analytiker ihre Arbeit tatsächlich auf solche Weise vor Kritik schützen. Fromm (1979) spricht daher von vulgärfreudianischen Immunsierungstendenzen.

Aufhebung der
Immunsierung

Was Freud selbst betrifft, so gilt ihm weder das "Ja" noch das "Nein" des Klienten per se als Beleg für die Richtigkeit der psychoanalytischen Deutung:

"Das direkte ‚Ja‘ des Analysierten ist vieldeutig. Es kann in der Tat anzeigen, dass er die vernommene Konstruktion als richtig anerkennt, es kann aber auch bedeutungslos sein oder selbst was wir ‚heuchlerisch‘ heißen können, indem es seinem Widerstand bequem ist, die nicht aufgedeckte Wahrheit durch eine solche Zustimmung weiterhin zu verbergen ...

Das ‚Nein‘ des Analysierten ist ebenso vieldeutig und eigentlich noch weniger verwendbar als sein ‚Ja‘. In seltenen Fällen erweist es sich als Ausdruck berechtigter Ablehnung; ungleich häufiger ist es Äußerung des Widerstandes ... Das ‚Nein‘ des Patienten beweist also nichts für die Richtigkeit der Konstruktion ... Die einzig sichere Deutung seines ‚Nein‘ ist ...

die auf Unvollständigkeit; die Konstruktion hat ihm gewiss nicht alles gesagt (Freud, 1937, 400).

Was die vulgärfreudianischen Immunisierungstendenzen betrifft, so können sie vermieden werden,

- indem an die Deutungen des Psychoanalytikers dieselben Wahrheitskriterien hinsichtlich methodischer Korrektheit, genetischer Argumentation, empirischer Reduktion und pragmatischer Evaluation angelegt werden wie an die des Klienten,
- und der Widerspruch des Klienten erst dann als Widerstand gilt, wenn der Klient entgegen einer *wahren Fremddeutung* an einer *falschen Selbstdeutung* festhält.

Indem Freud den Widerstandsbegriff (ähnlich wie Konrad Lorenz das Konzept des Aggressionstriebes) an bestimmten empirischen Konsequenzen festmacht, die sich aber erst ergeben, wenn der Klient seinen Widerstand aufgibt, liegt die pragmatische Evaluation der Fremddeutungen z.T. in der Zukunft. Daher kann (zumindest vorübergehend) die Situation eintreten, dass sowohl die Selbstdeutungen des Klienten als auch die Fremddeutungen des Psychoanalytikers Bestand haben, ohne dass eine Entscheidung herbeigeführt werden kann.

- So lange dies der Fall ist, können die Fremddeutungen des Analytikers daher nur *Vorschlagscharakter* beanspruchen, und das Festhalten des Klienten an seinen Selbstdeutungen kann nicht als Widerstand klassifiziert werden.
- Erst wenn der Klient die Fremddeutungen des Psychoanalytikers als neues Selbstverständnis übernimmt und sich die entsprechenden Konsequenzen einstellen, kann *rückblickend* davon gesprochen werden, dass dem Klienten seine bisher unbewussten Motive nunmehr bewusst geworden wären.

In diesem Sinne schreibt auch Freud (1937, 402):

"Wir geben die einzelne Konstruktion für nichts anderes aus als für eine Vermutung, die auf Prüfung, Bestätigung oder Verwerfung wartet."

Da Freud einen sehr weiten Widerstandsbegriff vertritt, erweist sich die Frage nach der Kritisierbarkeit der Psychoanalyse allerdings als weit komplexeres Problem, über das die Meinungen bis heute auseinandergehen (z.B. Therstappen, 1980; Grünbaum, 1988; Straßmaier, 2003). An einer Stelle spricht Freud sogar davon:

Die Komplexität des Kritisierbarkeitsproblems

"Was immer die Fortsetzung der Arbeit [der Psychoanalyse, W.K.] stört, ist ein Widerstand" (Freud, 1900, 495).

Der vulgärfreudianische Umgang mit dem "Ja" und dem "Nein" des Analysierten stellt angesichts eines so umfassenden Verständnisses von Widerstand nur die Spitze eines Eisberges dar.

Zudem unterscheidet Freud verschiedene Widerstandsformen, namentlich den Verdrängungs- und den Übertragungswiderstand, den Über-Ich- und den Es-Widerstand sowie den Widerstand aufgrund des sekundären Krankheitsgewinns (vgl. Thomä & Kächele, 1985) und betont, dass nicht die Bewussterdung verdrängter Inhalte per se therapeutische Wirkung entfalte, sondern allein die Überwindung der Widerstände (vgl. Mertens, 1991). Da jede Verhaltensweise zum Widerstand werden könne, komme es bei der Beurteilung, ob ein Widerstand vorliege, auf die individuellen Gegebenheiten des Einzelfalles an. Was beim einen ein Widerstand sei (z.B. Schweigen), könne beim anderen eine beginnende Individuation und Auflösung einer Übertragung sein.

Kontrollfragen

- Erläutern Sie die grundlegenden Unterschiede zwischen Beobachtungsprädikatoren und Deutungsprädikatoren!
- Wie lautet das intentionale Erklärungsmodell?
- Inwiefern nimmt die Psychoanalyse unter den intentionalen Erklärungen eine Sonderrolle ein?
- Erläutern Sie, was man unter Sinngehalten versteht!
- Erläutern Sie, worin das methodologische Hauptproblem der Psychoanalyse besteht!
- Erläutern Sie, was man in der Psychoanalyse unter Widerstand versteht!
- Erläutern Sie, inwiefern Gegenstandsverständnis und Wissensideal der Psychologie mit der Einführung des intentionalen Erklärungsmodells eine wesentliche Erweiterung erfahren!
- Erläutern Sie anhand eines Beispiels den Unterschied zwischen Bedingungs- und Begründungszusammenhängen!
- Zeigen Sie anhand eines Beispiels, dass es sich bei den nach dem Modell der intentionalen Erklärung konstruierten Begründungszusammenhängen nicht um empirische, sondern um strukturelle Gesetzmäßigkeiten handelt!
- Diskutieren Sie anhand eines Beispiels, ob bzw. inwiefern das intentionale Erklärungsmodell empirische Forschung überflüssig macht!
- Welche grundlegenden Unterschiede bestehen zwischen dem intentionalen und dem deduktiv-nomologischen Erklärungsmodell?
- Worauf beruht der im intentionalen Erklärungsmodell dargestellte strukturelle Zusammenhang zwischen Zielen und Mitteln?

- Warum haben Aussagen über Sinngehalte stets die Form von Hypothesen?
- Geben Sie eine schematische Darstellung von Handlungsabläufen und erläutern Sie, welche Anteile objektseitig und welche subjektseitig definiert sind!
- Was bedeutet die Interpretationsoffenheit von Handlungen und Sinngehalten, und woraus ergibt sie sich?
- Diskutieren Sie, ob bzw. unter welchen Voraussetzungen intentionale Erklärungen mit dem Anspruch der Erfahrungswissenschaftlichkeit verträglich sind:
 - a) Welche Haltung hat der Behaviorismus zu dieser Frage eingenommen?
 - b) Warum gilt die Auffassung des Behaviorismus inzwischen überholt?
 - c) Hat die Psychologie damit ihre Erfahrungswissenschaftlichkeit aufgegeben?
 - d) Wenn nein: warum nicht?
- Erläutern Sie den Terminus ‚handlungsauslösende Konstellation‘!
- Können intentionale Erklärungen an der Erfahrung scheitern? Wenn ja: wodurch?
- Was bedeuten in diesem Zusammenhang die Termini
 - a) ‚genetische Argumentation‘,
 - b) ‚empirische Reduktion‘ und
 - c) ‚pragmatische Evaluation‘?
- Erläutern Sie, wie eine intentionale Erklärung durch das Scheitern
 - a) der genetischen Argumentation
 - b) der empirischen Reduktion, bzw.
 - c) der pragmatischen Evaluationfalsifiziert werden kann! Geben Sie Beispiele dafür!
- Worin besteht der Unterschied zwischen Selbstdeutungen und Fremddeutungen?
- Welche Gründe kann es haben, wenn Selbstdeutungen und Fremddeutungen nicht übereinstimmen?
- Was versteht man unter dem Primat der Selbstdeutungen, wie ist es begründet, und inwiefern können die Selbstdeutungen dennoch widerlegt werden?
- Wie funktioniert die vulgärfreudianische Immunisierung psychoanalytischer Deutungen, und wie kann sie aufgehoben werden?
- Erläutern Sie, warum die Fremddeutungen des Analytikers zunächst nur Vorschlagscharakter beanspruchen können und welche Position Freud selbst diesbezüglich eingenommen hat!

3.3

Das narrative Erklärungsmodell

3.3.1 Reflexionstermini

Kontextabhängigkeit
von Bedeutungen

Aufgrund der Kontextabhängigkeit von Bedeutungen wäre es ein Missverständnis, wollte man die Genese der Sinngehalte nach dem Ursache-Wirkungs-Modell rekonstruieren. Schon die Worte, mit denen wir Einstellungen, Werte und Interessen beschreiben, auf deren Folie sich die Situation konstituiert, in der wir handeln, sind einer terminologischen Präzisierung nicht in dem Maße zugänglich, wie dies für eine deduktiv-nomologische Erklärung erforderlich wäre.

So weisen z.B. Kochinka & Werbik (2000) darauf hin, dass es keine Terminologie des Glücks gibt, die kontextfrei definiert, was wir unter Glück zu verstehen haben. Mit Worten wie Ehre, Treue, Rassismus oder Gerechtigkeit steht es nicht anders.

Geschichten und
Metaphern

Wenn wir jemanden fragen, was für ihn Glück bedeutet, werden nur die wenigsten eine operationale Definition geben, wie z.B. "Sex and Drugs and Rock'n'Roll" oder "Money for nothing and chicks for free". Sie werden eher Geschichten erzählen von Situationen, in denen sie Glück erlebt haben, oder sie werden dieses Glücksgefühl mit Metaphern umschreiben. Was das bedeutet, muss man aus dem Kontext verstehen, und dabei stellt sich dann ggf. heraus, dass selbst so coole Antworten wie die oben zitierten denn doch nicht als operationale Definition gemeint waren, sondern metaphorisch zu verstehen sind.

Reflexionstermini

Die Worte, mit denen wir Werte, Einstellungen und Interessen beschreiben, sind nicht eindeutig, sondern haben ein diffuses, an den Rändern offenes Bedeutungsfeld, das vom Außenstandpunkt des Forschers nur mittels hermeneutischer Analyse erhellt werden kann bzw. vom Innenstandpunkt des handelnden Subjekts der Reflexion bedarf. Weil wir in den Sozialwissenschaften ohne solche Worte nicht auskommen, stellen sie als Reflexionstermini

dennoch einen unverzichtbaren Bestandteil der psychologischen Wissenschaftssprache dar.

3.3.2 Die narrative Erklärung

Auch die Erklärungsaufgabe, der wir uns gegenübersehen, wenn wir die Genese von Sinngehalten zu rekonstruieren versuchen, ist eine gänzlich andere als im deduktiv-nomologischen Erklärungsmodell. Sie gleicht in ihrer Struktur jener Art von Erklärungen, welche wir geben, indem wir eine Geschichte erzählen, und die Danto (1974) zum narrativen Erklärungsmodell ausgearbeitet hat.

Erklärung als
Präzisierung des
Explanandums

Wenn wir danach fragen, warum jemand unter den objektseitig gegebenen Bedingungen zu bestimmten Sinngehalten (und entsprechenden Handlungen) gekommen ist – z.B., warum er nach so vielen Ehejahren wieder allein leben wollte (oder warum er seine Frau verlassen hat) –, dann fragen wir nach einer *Präzisierung des Explanandums*. Dabei ist nicht irgendein Sachverhalt zu explizieren, sondern eine *Veränderung* (etwa: "Du warst doch immer so ein Familienmensch, und jetzt willst du allein leben?") bzw. eine *Transformation*, welche die objektseitig gegebenen Bedingungen in die jeweiligen Sinnhalte und Handlungen übersetzt (etwa: "In deiner Ehe hast du doch alles gehabt, was man sich wünschen kann, und jetzt willst du das alles aufgeben?").

Das Explanandum besteht also in der Diskrepanz zwischen den aus zwei Perspektiven (früher vs. jetzt bzw. objektseitig vs. subjektseitig) bestehenden Sachverhalten, die durch das Explanans miteinander in Beziehung gesetzt werden müssen. Eine solche sinnstiftende Verbindung der beiden perspektivischen Eckpunkte leistet das Erzählen einer Geschichte, indem sie durch ihre Darstellung erklärt, *wie* die Veränderung bzw. Transformation stattgefunden hat.

Straub (1999, 274) weist in diesem Zusammenhang darauf hin, dass sich Beschreibungen und Erklärungen mitunter nicht fein säuberlich voneinander trennen lassen. So sei z.B. das Konzept der dichten Beschreibung nach Geertz (1987)

Beschreiben und
Erklären

"... genau besehen, nicht bloß an einen bestimmten Begriff der Beschreibung gekoppelt. Dichte Beschreibungen machen uno actu Verständnis- und Erklärungsangebote – sie liefern Erklärungen, die zeigen, was genau in Frage steht und worum es dem Fragenden und/oder Antwortenden eigentlich geht, Antworten, die schließlich klarmachen, wie es möglich war

und dazu kam, dass das fragliche Ereignis eintrat oder unterblieb, eine sinn- und bedeutungsstrukturierte Handlung vollzogen oder unterlassen wurde."

Und Schwemmer (1987) weist darauf hin, dass eine Handlungsbeschreibung, die die Bedeutung des Handelns nicht übergehen will, niemals lediglich die Beschreibung einer isolierten Tätigkeit, sondern immer eine Gesamtbeschreibung der Handlung und ihres Kontextes ist.

Damit werden die Vorgeschichte, der Verlauf und die (erwartete) Nachgeschichte einer Handlung zentral, wobei die Temporalstruktur des Handelns dadurch explizit wird, dass die Geschichten, in die das fragliche Handeln eingebettet ist, und die Geschichte, die es möglicherweise selbst verkörpert, erzählt werden:

"Sobald eine Handlung als temporal komplexes Problem aufgefasst wird, was der Fall ist, wenn sie als Endpunkt einer erzählten Geschichte präsentiert wird, wird ein Anfang und ein Verlauf einer Geschichte vorausgesetzt, die plausibilisieren, wie es zu der fraglichen Handlung hat kommen können, wo diese doch in der Vergangenheit – einem Zeitraum, den die Erzählung konkretisiert – noch nicht intendiert, nicht erwartet und womöglich nicht einmal vorstellbar war." (Straub, 1999, 278)

Schematische
Darstellung der
narrativen Erklärung

Ähnlich wie die oben behandelten Erklärungsmodelle kann auch das Modell der narrativen Erklärung schematisch dargestellt werden:

- (1) $x \in F$ in t_1
- (2) H ereignet sich mit x in t_2
- (3) $x \in G$ in t_3

Zu einem Zeitpunkt t_1 bzw. aus objektseitiger Perspektive besteht der Sachverhalt $x \in F$. Zum Zeitpunkt t_3 bzw. aus subjektseitiger Perspektive besteht der Sachverhalt $x \in G$.

Diese Veränderung bzw. Transformation von $x \in F$ in t_1 nach $x \in G$ in t_3 wird durch ein dazwischenliegendes Ereignis H von beliebigem Komplexitätsgrad *erklärt*, indem dieses eine *sinnstiftende Verbindung* zwischen den Sachverhalten $x \in F$ in t_1 und $x \in G$ in t_3 herstellt.

Die Sätze (1) und (3) bilden zusammen das Explanandum; (2) ist das Explanans. Die Hinzuziehung von (2) ergibt die Erklärung für (1) \rightarrow (3), womit die Erklärung die Form einer Erzählung annimmt. "Sie hat einen Anfang (1), einen Mittelteil (2) und ein Ende (3)" (Danto, 1974, 376).

3.3.3 Sozialwissenschaftliche Hermeneutik

Anders als die deduktiv-nomologische Erklärung, die auf die Frage nach den *Ursachen* antwortet, *warum* ein Sachverhalt eingetreten ist, und anders als die intentionale Erklärung, die auf die *Ziele* und damit auf das *Wozu* von Handlungen gerichtet ist, fragt die narrative Erklärung nach dem *Verlauf*, d.h. danach, *wie* eine Veränderung bzw. Transformation stattgefunden hat. Auch dabei lassen sich ggf. typische Verlaufsformen oder -gesetzmäßigkeiten identifizieren, deren Erforschung jedoch dem Paradigma der qualitativen Sozialforschung folgen muss, das an gänzlich andere methodologische Voraussetzungen gebunden ist als etwa die Hypothesenprüfung im quantitativen Paradigma, das sich an der deduktiv-nomologischen bzw. statistisch-induktiven Erklärung orientiert.

Qualitative
Sozialforschung

Ausgangspunkt der qualitativen Sozialforschung ist es, dass die subjektive Welt des Menschen als eine sinn- und bedeutungsstrukturierte Wirklichkeit begriffen wird. Die Bedeutungen, die mit dem Denken, Sprechen, Handeln und dessen Objektivationen verbunden sind, lassen sich jedoch nicht distanziert beobachten, sondern das Erkennen sinn- und bedeutungsstrukturierter Wirklichkeiten stützt sich notwendigerweise auf Deutungs- und Interpretationsleistungen und ist somit auf sinnverstehende Methoden angewiesen, bei denen es zudem nicht – oder zumindest nicht primär – darum geht, bestimmte Erfahrungen unter möglichst standardisierten, kontrollierten und reproduzierbaren Bedingungen herzustellen, sondern darum, die ohnedies gemachten lebensweltlichen Erfahrungen, Handlungen und Interaktionen in ihrem Sinnzusammenhang zu verstehen. Diese Orientierung an der Praxis sowie an den Selbst- und Weltverständnissen von Menschen bezeichnen Appelsmeyer, Kochinka & Straub (2000, 712) als das erfahrungswissenschaftliche Prinzip qualitativer Forschung und weisen zugleich darauf hin, dass dem qualitativen Paradigma auch ein anderer Erfahrungsbegriff zugrunde liegt, der dem ursprünglichen, aristotelischen Sinn des Wortes *empeiria* sehr viel näher kommt als der Erfahrungsbegriff des quantitativen Paradigmas, der auf eine methodisch-technisch produzierte Empirie abzielt. Dies hat weitreichende methodologische Konsequenzen.

Im quantitativen Paradigma steht der Bestätigungsgrad empirischer Hypothesen im Vordergrund des methodischen Kontrollinteresses. Insofern kann von einem Primat der Theorie vor der

Methodologische
Konsequenzen

Empirie gesprochen werden: Die Theorie entscheidet darüber, welche Ausschnitte aus der empirischen Wirklichkeit als relevant erachtet werden. Im qualitativen Paradigma besteht dagegen ein umgekehrter Primat, wie ihn etwa Hoffmann-Riem (1980) im Verzicht auf eine Hypothesenbildung ex ante formuliert hat und wie er auch in der von Blumer (1973) getroffenen Unterscheidung zwischen explorativer und inspektiver Forschung zum Tragen kommt.

Konkret bedeutet dies zunächst, dass den Mitgliedern des Forschungsfeldes, welches untersucht wird, der Expertenstatus bezüglich ihrer eigenen subjektiven Wirklichkeit zuzuerkennen ist (Schütze, 1977). Zweitens ergibt sich daraus die Bevorzugung offener Erhebungsinstrumente, bei denen die Forschungssubjekte nicht von vornherein auf bestimmte Antwort- oder Reaktionsmöglichkeiten festgelegt sind. Drittens führt dies zu ganz anderen Untersuchungsmethoden und Techniken, für die z.B. Messung, Mathematisierung und Experiment keine zentrale Rolle spielen; und viertens ist auch die Objektivität der Forschungsergebnisse in der qualitativen Sozialforschung auf anderem Wege zu gewährleisten: nicht durch Distanzierung des Forschers vom Forschungsgegenstand, sondern durch Transzendierung seiner Subjektivität.

Insofern, als diese methodologischen Konsequenzen in einem diametralen Gegensatz zu jenen stehen, die sich im quantitativen Paradigma ergeben, ist der Antagonismus, mit dem die Vertreter der beiden Paradigmen einander lange Zeit begegnet sind, nicht weiter verwunderlich. Zumindest in der Soziologie hat sich inzwischen aber die Auffassung durchzusetzen begonnen, dass quantitative und qualitative Methoden keinen Gegensatz darstellen, sondern einander fruchtbar ergänzen (vgl. Kelle & Erzberger, 2000).

Kontrollfragen

- Nennen Sie Gründe, warum die Genese von Sinngehalten nicht nach dem Ursache-Wirkungs-Modell rekonstruiert werden kann!
- Erläutern Sie, was man unter Reflexionstermini versteht und wodurch sich diese von den Beschreibungstermini unterscheiden!
- Erläutern Sie, warum sich Beschreibung und Erklärung von Handlungen mitunter nicht voneinander trennen lassen!
- Erläutern Sie, was Temporalstruktur des Handelns bedeutet!

- Stellen Sie das narrative Erklärungsmodell schematisch dar und erläutern Sie, worin dabei
 - a) das Explanandum sowie
 - b) das Explanans besteht!
- Erläutern Sie, inwiefern die narrative Erklärung die Form einer Erzählung hat!
- Wie unterscheiden sich
 - a) die deduktiv-nomologische,
 - b) die intentionale und
 - c) die narrative Erklärung hinsichtlich ihrer Fragestellung?
- Worin besteht der Ausgangspunkt der qualitativen Sozialforschung?
- Warum ist die qualitative Sozialforschung auf sinnverstehende Methoden angewiesen?
- Worin besteht das erfahrungswissenschaftliche Prinzip qualitativer Forschung (im Unterschied zur quantitativen Forschung)?
- Welches sind die methodologischen Konsequenzen des qualitativen Paradigmas?

3.4

Sinnrationalität

3.4.1 Regelgeleitetes Handeln

Handlungsergebnis
und Handlungsfolgen

Mit der Einführung von Reflexionstermini, wie wir sie z.B. verwenden, um den Inhalt von Einstellungen und Werten darzustellen, sind wir über das intentionale Erklärungsmodell in der Form, wie wir es bisher kennengelernt haben, insofern bereits hinausgegangen, als wir bisher nur eine bestimmte Form von finalen Handlungsorientierungen (Zielen) kennengelernt haben: die *Zwecke*.

Beispielsweise schon bei der in der Sozialisation gewaltbereiter Jugendlicher vermittelten Norm, in der Auseinandersetzung mit anderen seine Fäuste zu gebrauchen (Bandura & Walters, 1959), handelt es sich aber offensichtlich nicht um einen Zweck, den zu verfolgen den Kindern von ihren Eltern vorgeschrieben wird, sondern es geht um eine *Regel*, wie sie sich im Streit mit Gleichaltrigen "richtig" zu verhalten haben, nämlich: sich nichts gefallen zu lassen und sich mit allen Mitteln durchzusetzen.

Der Jugendliche kann kämpfen, weil er meint, damit seine Ziele durchsetzen zu können, und/oder weil er meint, dass es richtig ist, zu kämpfen (selbst wenn er keine Chance hat, seine Ziele damit zu erreichen).

Wir müssen daher unterscheiden zwischen:

- Zwecken, die sich auf die Folgen einer Handlung beziehen, und
- Handlungsregeln, die sich auf das Ergebnis der Handlung beziehen.

Als Handlungsergebnis bezeichnet man – im Unterschied zu den Folgen einer Handlung – nach von Wright (1974) jenen Zustand, der mit Beendigung einer Handlung besteht und ohne den die Handlung nicht vollzogen ist.

Handlungsregeln

Handlungsregeln (zu denen u.a. auch die Vorsätze gehören) legen weder das Wie noch das Wozu einer Handlung fest, sondern

lediglich, *dass* die Handlung (in bestimmten Situationen) vollzogen werden muss.

Werden solche Regeln allgemein befolgt, so braucht man neben dem Hinweis auf die Regel keine weiteren Gründe anzugeben, um die Handlung zu erklären. Warum man z.B. geantwortet hat, wenn man gefragt wurde, bedarf keiner weiteren Erklärung (außer eben dem Hinweis auf die Regel, dass man auf Fragen antwortet).

Dass es sich auch dabei um eine finale Handlungsorientierung handelt, liegt auf der Hand: Wenn man geantwortet hat, so ist die Handlung damit abgeschlossen.

3.4.2 Nonfinale Handlungsorientierungen

Afinale Aufforderungen zur Ausführung einer Handlung können aber nicht nur mit finalen Aufforderungen zur Erreichung eines bestimmten Zieles verbunden sein,

Finale und nonfinale Handlungsorientierungen

- welches ein bestimmtes Handlungsergebnis (regelgeleitetes Handeln)
- oder bestimmte Handlungsfolgen (zweckgebundenes Handeln)

zum Inhalt hat, sondern sie können auch mit nonfinalen Aufforderungen verbunden werden, die sich weder auf das Handlungsergebnis noch auf die Handlungsfolgen, sondern auf die Form des Handelns beziehen.

Gerechtigkeit oder Fairness, aber auch Spaß oder Glück sind Werte, an denen wir ein *Interesse* haben, die – anders als die Handlungsziele – nicht einfach entfallen, wenn wir sie einmal verwirklicht haben, sondern als Handlungsorientierungen weiter bestehen bleiben, da sie nicht einzelne Handlungen in unserem Leben, sondern den *Sinnzusammenhang* unseres Handelns betreffen.

Innerhalb der nonfinalen Handlungsorientierungen kann unterschieden werden zwischen Interessen im engeren und im weiteren Sinne.

Interessen im engeren Sinne (z.B. Wahrheit, Zufälligkeit) werden durch Beurteilungstermini beschrieben und sind durch klar explizierbare Herstellungsprinzipien definiert.

Interessen im engeren Sinne

- Im Falle der Wahrheit sind dies die Regeln der Geltungsprüfung.
- Im Falle der Zufälligkeit sind es die Konstruktionsprinzipien für Zufallsgeneratoren (Eindeutigkeit, Ununterscheidbarkeit, Wiederholbarkeit).

Diese Prinzipien definieren gleichsam *Ideale*, und die Interessen i.e.S. sind insofern *nonfinal*, als diese Ideale niemals definitiv erfüllt sein können.

- Eine universelle empirische Hypothese lässt sich nicht beweisen, weil wir immer nur endlich viele Beobachtungen haben.
- Auch der beste Würfel (Regelmäßigkeit der Form, Homogenität und Härte des Materials) erfüllt die Konstruktionsprinzipien für Zufallsgeneratoren nur *cum grano salis*, und wenn man die Stichprobe hinreichend groß macht, kann bei zweiseitiger Fragestellung jede Nullhypothese verworfen werden.

Trotzdem kommen Handlungssequenzen, die von Interessen i.e.S. geleitet sind, irgendwann zu einem Abschluss und sind insofern *quasi-final*:

- Früher oder später kommen wir zu der Überzeugung, eine Hypothese habe sich hinreichend bewährt, so dass wir sie zu unserem gesicherten Wissen zählen können, das keiner weiteren Überprüfung bedarf.
- Früher oder später kommen wir zu der Überzeugung, unser Würfel sei hinreichend regelmäßig, homogen und hart, so dass wir ihn praktisch anwenden können und es keiner weiteren Verbesserung des Würfels bedarf.

Interessen im weiteren Sinne

Schon bei Werten wie Gerechtigkeit oder Fairness liegen die Dinge deutlich anders. Zwar können wir durch bestimmte Regeln (Gesetzbücher) definieren, was Recht oder Unrecht ist, aber Gerechtigkeit ist mehr und anderes als bloß buchstabengetreue Gesetzesbefolgung. Gerechtigkeit ist auch nicht quasi-final. Wenn es uns einmal gelungen ist, eine einigermaßen gerechte Praxis zu etablieren, bleibt das Interesse an der Gerechtigkeit trotzdem weiter bestehen.

Gerechtigkeit ist damit ein Beispiel für ein Interesse im weiteren Sinne, das nicht nur einzelnen Handlungen (z.B. Gerichtsurteilen) oder Handlungssequenzen (z.B. dem Prozessverfahren) einen Sinn verleiht, sondern dem gesamten Leben eines jeden, für den

Gerechtigkeit einen Wert darstellt, und das als Orientierung seines Handelns permanent bestehen bleibt.

Interessen i.w.S. geben dem Leben eine sinnhafte Form und können daher mit Kambartel (1981) auch als Lebensorientierungen bezeichnet werden.

Lebensorientierungen

- Lebensorientierungen sind im Unterschied zu den Handlungsorientierungen i.e.S. nichts, wozu man sich explizit auffordert (z.B. "Sei glücklich!"), obwohl man das auch tun kann.
- Lebensorientierungen dienen vielmehr als Interpretationsfolie, unter der das, was wir erleben und tun, in einen sinnhaften Zusammenhang gebracht wird.

Insofern sind Lebensorientierungen maßgeblich an der Konstitution der Situation und der handlungsauslösenden Konstellation beteiligt, *in welcher* wir diese oder jene Handlungsorientierungen i.e.S. verfolgen.

- Die Interpretation der (objektseitig definierten) Umweltbedingungen, durch welche sich die (subjektseitig definierte) Situation konstituiert, erfolgt aber in der Regel *nicht bewusst*, sondern die subjektive Wirklichkeit, in der wir leben und handeln, erscheint uns als *selbstverständlich*.
- Lebensorientierungen macht man sich daher meist erst dann bewusst, wenn man sich in einer Sinnkrise befindet, d.h. wenn einem der Sinnzusammenhang seines Handelns und Lebens eben nicht mehr selbstverständlich ist.

Den Prozess, in dem man sich des Sinnzusammenhanges bewusst zu werden versucht, bezeichnet man als Reflexion.

3.4.3 Exkurs: Die Perspektivität des Gewaltbegriffs

Anders als die Interessen i.e.S. können Lebensorientierungen nicht durch klar explizierbare Herstellungsprinzipien definiert werden, sondern werden durch Reflexionstermini beschrieben, die – wie wir bereits gesehen haben – ein an den Rändern offenes Bedeutungsfeld haben.

Beispiel
Gewaltfreiheit

Als Beispiel betrachten wir die Orientierung der Gewaltfreiheit, auf die wir bei der Bestimmung der Aufgaben der Friedenswissenschaften zurückgegriffen hatten.

- Dass es sich dabei um eine nonfinale Orientierung handelt,

liegt auf der Hand: Wenn es jemandem einmal gelungen ist, erfolgreich gewaltfreien Widerstand zu leisten, so entfällt das Motiv der Gewaltfreiheit damit nicht, sondern bleibt weiterhin ein Wert, an dem er sich auch bei künftigen Widerstandshandlungen orientieren wird (es sei denn, dass er sich – nachdem er eine Sinnkrise durchgemacht hat – von der Gewaltfreiheit verabschiedet).

- Auch wird sofort deutlich, dass der Inhalt dieser Orientierung ein an den Rändern offenes Bedeutungsfeld hat, wenn man sich der Diskussion erinnert, die gegen Ende der 80er Jahre des letzten Jahrhunderts um die These entbrannt ist: "Gewaltfreier Widerstand ist Gewalt".

Personale und strukturelle Gewalt

Im Unterschied zu Aggression ist Gewalt ein *normativer Begriff*, wobei wir in Anschluss an Galtung (1969) bereits oben zwischen intendierter und nicht intendierter sowie zwischen personaler (oder direkter) und struktureller (indirekter) Gewalt unterschieden haben.

Wesentlich am Begriff der *strukturellen Gewalt* ist, dass es hier keinen Gewalttäter gibt; und schon gar keinen, der die Gewalt mit Absicht ausübt. Dennoch kann strukturelle Gewalt nicht minder tödlich sein.¹

Strukturelle Gewalt ist nach Galtung (1969, 13) in das System eingebaut und äußert sich in ungleichen Machtverhältnissen und folglich in ungleichen Lebenschancen. So schreibt Galtung z.B., dass "in einer Gesellschaft, in der die Lebenserwartung der Oberschicht doppelt so hoch ist wie die der Unterschicht, Gewalt manifest [ist], auch wenn keine konkreten Akteure sichtbar sind, die direkt gegen andere vorgehen, wie etwa im Fall der Tötung eines Menschen durch einen anderen."

Positiver Friedensbegriff

Eines der zentralen Motive für die Erweiterung des Gewaltbegriffs über personale Gewalt hinaus resultiert aus der Einsicht, dass (auch verdeckte) Gewalt zu Gegengewalt führt.

Wenn z.B. die ökonomischen Verhältnisse dazu führen, dass ein Teil der Bevölkerung seine Grundbedürfnisse nur noch auf illegalem Weg befriedigen kann, durch Raub und Diebstahl, wie etwa die Straßenkinder in Brasilien, dann sind diese Straßenkinder zwar

1. So ist z.B. der Obdachlose, der in den Strassen von Berlin erfriert, Opfer struktureller Gewalt. Der Obdachlose, der von Skinheads erschlagen wird, fällt dagegen personeller Gewalt zum Opfer.

einerseits Gewalttäter, d.h. Urheber personaler Gewalt, aber sie sind dies deshalb, weil sie selbst Opfer der Verhältnisse sind, welche sie dazu zwingen, sich zum bloßen Überleben gewaltförmiger Mittel zu bedienen.

Entsprechend verfolgt die Kenntlichmachung struktureller Gewalt bei Galtung vor allem das Ziel, auch die verdeckten Gewaltverhältnisse zu bekämpfen und einen positiven Friedensbegriff zu begründen, der Frieden nicht nur als De-facto-Abwesenheit direkter personaler Gewalt begreift, sondern als einen Zustand, in dem auch niemand gezwungen ist, zwischen den Alternativen der Gewaltanwendung oder des eigenen Unterganges zu wählen.

Obwohl die Galtung'sche Gewaltdefinition natürlich nicht unwidersprochen blieb, erfreute sie sich sowohl in der sozialwissenschaftlichen Friedensforschung als auch in der Friedensbewegung der ausgehenden 70er und beginnenden 80er Jahre eines großen Einflusses. Dieser war so groß, dass z.B. die baden-württembergische Landesregierung meinte, ihre Landeskinder davor schützen zu müssen: Am Höhepunkt des Widerstandes gegen den NATO-Doppelbeschluss wurde die Verwendung des Begriffes der strukturellen Gewalt an baden-württembergischen Schulen vom damaligen Kultusminister Mayer-Vorfelder per Erlass verboten.

Skandalisierung und
Diskreditierung

Der Grund hierfür ist wohl darin zu suchen, dass der Begriff der strukturellen Gewalt – um es mit den Worten von Albert Fuchs (1993) auszudrücken – "bestens geeignet" erscheint, "alle Verhältnisse zu skandalisieren, die Opfer hervorbringen, und ihre Repräsentanten zu diskreditieren". Die so Diskreditierten wehren sich dann mit Mitteln, die an George Orwells Roman *1984* erinnern – durch terminologische Säuberung.

Die Skandalisierung bestehender Gewaltverhältnisse dürfte durchaus in Galtungs Sinne sein, die Diskreditierung ihrer Repräsentanten dagegen weniger.

Tatsächlich trägt der Begriff der strukturellen Gewalt ja die Chance in sich, Gewaltverhältnisse bekämpfen zu können, ohne dass jemand dafür persönlich verantwortlich zu machen ist. Andererseits begünstigt das übliche Vorverständnis von Gewalt als personaler Gewalt jedoch sowohl die Tendenz, tatsächliche oder vermeintliche Repräsentanten der bestehenden Verhältnisse persönlich anzuklagen, sowie deren Neigung, sich durch die Skanda-

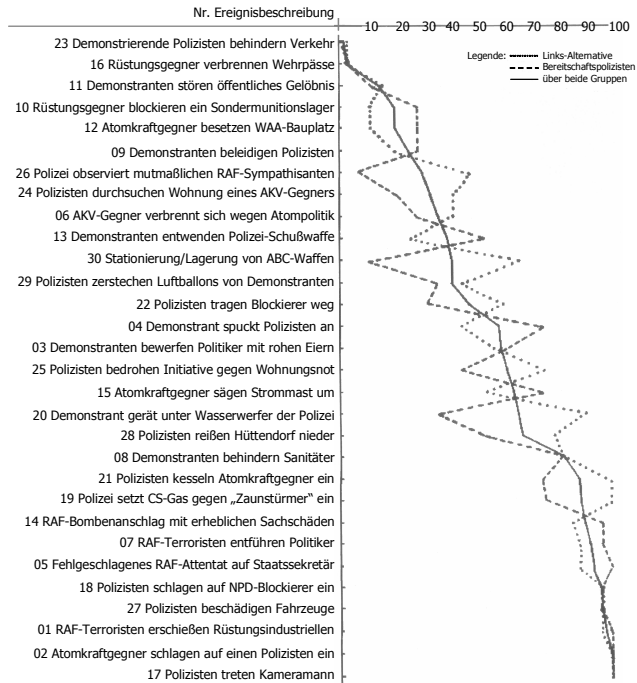
lisierung bestehender Gewaltverhältnisse persönlich angegriffen zu fühlen.

Unschärfe des Gewaltbegriffs

Das Problem der Galtung'schen Gewalttypologie liegt entsprechend weniger in der Unterscheidung zwischen personaler und struktureller Gewalt – die für die Kenntlichmachung verdeckter Gewalt grundlegend ist – als in der Unschärfe des Gewaltbegriffs, namentlich in seinem extrem negativen Wertungsgehalt und in der perspektivischen Verwendungsweise des Gewaltbegriffs, welche für die Alltagspraxis typisch ist.

So fanden Fuchs & Maihöfer (1992), dass zeitgeschichtliche Protestaktivitäten und Repressionsmaßnahmen von Mitgliedern links-alternativer Gruppen und Auszubildenden der Bereitschaftspolizei sehr unterschiedlich bewertet wurden (vgl. Abb. 3.4.1).

Abb. 3.4.1: Beurteilung von Protestaktivitäten und Repressionsmaßnahmen als "ziemlich ..." bzw. "absolut sicher Gewalt" (nach Fuchs, 1993)



Während in den Extrembereichen kein nennenswerter Unterschied zwischen den beiden Informantengruppen bestand, legten im Unsicherheitsbereich fast durchgehend

- die Polizisten bei der Beurteilung von Protestaktivitäten und

- die Links-Alternativen bei der Beurteilung von Repressionsmaßnahmen

einen weiteren Gewaltbegriff zugrunde.

Fuchs sieht darin ein Indiz für einen einheitlichen Begriffskern bei gleichzeitiger perspektivischer Verwendungsweise des Gewaltbegriffs (vgl. Abb. 3.4.2). Während die Gewaltförmigkeit eigenen Verhaltens systematisch unterschätzt wird, wird die des gegnerischen Verhaltens systematisch überschätzt.

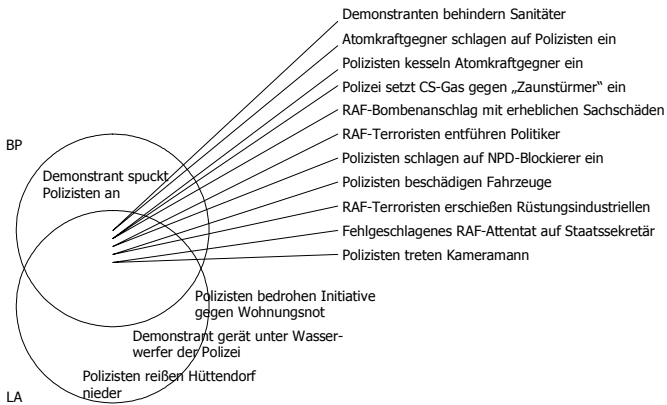


Abb. 3.4.2: Der Begriffskern des Gewaltbegriffs (nach Fuchs, 1993)

Die Qualifizierung des jeweiligen gegnerischen Verhaltens als Gewalt verspricht u.a. strategische Vorteile. Aufgrund der negativen Valenz des Gewaltbegriffs ermöglicht sie es, den Akteur gegenüber Dritten zu diskreditieren und zu diffamieren und dadurch seinen Handlungsspielraum einzuschränken, aber auch gleichzeitig eigene Gegengewalt präventiv zu rechtfertigen.

Vergeltungsmoral

Umgekehrt sichert oder eröffnet die Beurteilung des eigenen Verhaltens als gewaltfrei den eigenen Handlungsspielraum und erleichtert dessen Rechtfertigung (Neidhardt, 1986).

Dass Akteur und Rezipient ein konkretes Handlungssegment in aller Regel konträr beurteilen, wird vor allem in Verbindung mit der verbreiteten Orientierung am Reziprozitätsprinzip (d.h. am Maßstab äquivalenter Gegenseitigkeit) kontraproduktiv. Denn dann ist etwa folgende Schlussfolgerung zu unterstellen:

Person X hat (grundlos) gegen mich (oder unbescholte andere) die Gewalttat Y begangen.

(Unprovozierte) Gewalt verdient (entsprechende) Gegengewalt.

Folglich bin ich (oder irgendjemand anderes) berechtigt, gegen X (entsprechende) Gegengewalt auszuüben.

Demnach droht die brisante Mischung von Perspektivendivergenz der Gewaltzuschreibung und Orientierung am Reziprozitätsprinzip eine destruktive Eskalationsdynamik nahezu zwangsläufig hervorzutreiben.

Um dem entgegenzuwirken, möchte Fuchs die Dehnbarkeit des Gewaltbegriffs möglichst einschränken.

Sprachnormierung
als Sackgasse

Selbst wenn man so weit geht wie Fuchs und nur noch die Verletzung der körperlichen Unversehrtheit als Gewalt gelten lässt, ist damit jedoch wenig gewonnen. Denn zum einen sind auch Verletzung und Unversehrtheit keine wertfreie Begriffe, so dass das Problem lediglich verschoben wird. Zum anderen kann die damit angestrebte Reduzierung des Gewaltbegriffs auf seinen Begriffskern allenfalls für den wissenschaftlichen Diskurs über Gewalt Verbindlichkeit erlangen, nicht jedoch für die politischen und alltagsweltlichen Legitimationsdiskurse, in denen Gewaltanwendung realiter gerechtfertigt wird. Menschen handeln aufgrund dessen, was sie ganz persönlich als Gewalt empfinden, und nicht aufgrund dessen, wie der Terminus von Wissenschaftlern gebraucht wird. Drittens kann auch die oben skizzierte Eskalationsdynamik nur insofern konkret erforscht werden, als die tatsächlichen Orientierungen der Akteure (namentlich die Unschärfe und Perspektivität ihres Gewaltverständnisses) rekonstruiert werden, so dass sich das Programm einer Normierung von Reflexionstermini als gleichermaßen illusorisch wie kontraindiziert erweist.

Fazit: Sobald Lebensorientierungen (z.B. Gewaltfreiheit) ins Blickfeld sozialwissenschaftlicher Forschung gelangen, genügt es nicht mehr, die Worte, mittels derer sie beschrieben werden, terminologisch zu normieren, sondern ihre Bedeutung muss aus der je subjektiven Perspektive der Probanden selbst rekonstruiert werden.

Dies tritt spätestens dann ein, wenn wir das psychologische Laboratorium verlassen und nicht nur isolierte Handlungen betrachten, sondern größere Handlungszusammenhänge. Während es im ersteren Fall noch angehen mag, die Geltungsprüfung auf die empirische Reduktion zu beschränken, ist die genetische Rekonstruktion der Handlungen im zweiten Fall unverzichtbar.

3.4.4 Kontextabhängigkeit

Man kann zwar versuchen, Werthaltungen oder Einstellungen mittels Fragebögen zu operationalisieren, aber auch dabei können die Antworten (und bereits die Fragen) nur in ihrem Kontext verstanden werden.

Kontextabhängigkeit

Wenn wir bei der Einstellungsmessung so tun, als wären die Antworten eindeutig und könnten einfach verrechnet werden, liegt dies entweder daran, dass wir das Problem der Kontextabhängigkeit übersehen haben (was einen Methodenfehler darstellt), oder dass wir die Fragen so gestellt haben, dass der Kontext eindeutig ist (und nicht weiter problematisiert zu werden braucht).

Wenn wir jemanden fragen "Wie geht es Ihrer Familie?", werden wir darauf ganz unterschiedliche Antworten erhalten, je nachdem, ob wir vorher gefragt haben "Wie geht es Ihrem Chef?" oder "Wie geht es ihrer Frau?" (Das lässt sich empirisch sehr schön demonstrieren).

Im ersten Fall werden wir uns wundern, wenn der Gefragte die Gattin in seiner Antwort mit keinem Wort erwähnt. Im zweiten Fall ist uns dieselbe Antwort selbstverständlich, da wir über das Wohlergehen der Gattin bereits in der Antwort auf die vorangegangene Frage unterrichtet wurden.

Wenn wir z.B. auf einer Fete gerade dazukommen, wie ein Freund (A) einen anderen (B) nach dem Wohlergehen seiner Familie fragt und nun mitanhören, wie die Gattin in der Antwort völlig unerwähnt bleibt, dann wird uns das – wie gesagt – wundern, und wir werden B vielleicht beiseite nehmen und fragen: "Sag einmal, was ist los mit eurer Beziehung, du hast doch früher immer so begeistert von deiner Frau erzählt, und jetzt tust du, als ob es deine Frau gar nicht gäbe".

Narrative
Bedeutungsklärung

Und B wird vielleicht antworten: "Die hat sich im Urlaub einen Liebhaber zugelegt, und wir haben uns getrennt", oder er wird la-

chen und sagen: "Wie kommst du denn auf die Idee, dass etwas nicht stimmen könnte, ich habe doch dem A gerade erst erzählt, was wir für einen wunderschönen Urlaub hatten und wie sie das genossen hat ..."

Tab.3.4.1: Narrative Bedeutungsklärung durch Rekonstruktion des Kontextes, in welchem B's Aussage zu verstehen ist

Fall 1	
t ₁ = früher	B erzählt begeistert von seiner Frau
t ₂ = im Urlaub	Seine Frau legt sich einen Liebhaber zu, und die beiden trennen sich
t ₃ = jetzt	B lässt seine Frau in der Antwort auf A's Frage nach dem Wohlergehen seiner Familie unerwähnt
Fall 2	
t ₁ = früher	B erzählt begeistert von seiner Frau
t ₂ = gerade vorher	B erzählt in Antwort auf A's Frage nach dem Wohlergehen seiner Frau begeistert vom gemeinsamen Urlaub
t ₃ = jetzt	B lässt seine Frau in der Antwort auf A's Frage nach dem Wohlergehen seiner Familie unerwähnt

Wie unschwer zu erkennen ist, folgt die Bedeutungsklärung dem narrativen Erklärungsmodell (vgl. Tab. 3.4.1). Mit dem Erzählen einer Geschichte rekonstruiert B den Kontext, in welchem seine Antwort auf A's Frage zu verstehen ist, und erklärt damit, warum er seine Frau unerwähnt ließ.

Methodisch kontrolliertes Verstehen

Während das oben diskutierte Beispiel zeigt, dass die Kontextabhängigkeit von Bedeutungen nicht erst bei der Rekonstruktion von Werten oder Einstellungen zum Methodenproblem wird, ist es nur ein Bruchteil der psychologischen Sachverhalte, welche deutungsfrei beschreibbar sind (vgl. Tab. 3.4.2).

Schon Wahrnehmungen und Emotionen können vom Außenstandpunkt des Forschers nicht einfach beobachtet, sondern müssen aus dem beobachtbaren Verhalten erschlossen werden. Daran ändert sich auch nichts, wenn wir auf Selbstauskünfte der Probanden zurückgreifen. Auch diese können nicht deutungsfrei verstanden werden, und die Frage des methodisch kontrollierten Verstehens erweist sich folglich als ein zentrales Methodenproblem der Psychologie schlechthin.

Termini		Beispiele	Überprüfung
Logische Partikel	Junktoren	Negation, Konjunktion, Adjunktion, Subjunktion	Regeln der Einführung sind zugleich die Regeln der Überprüfung
	Quantoren	Allquantor, Einsquantor	
Beschreibungs- termini	Objektseitig definierte Beschreibungstermini (= Beobachtungstermini)	Stimuli und Verhalten	Zur Überprüfung bedarf es der Beobachtung
	Subjektseitig definierte Beschreibungstermini	Wahrnehmungen und Emotionen	Zur Überprüfung bedarf es je nach Perspektive der Selbstbeobachtung oder der Beobachtung und der Deutung
	Handlungstermini	Handlungen und Handlungszusammenhänge	Zur Überprüfung bedarf es der Beobachtung und der Deutung
Deutungstermini	Deutungstermini i.e.S.	Zwecke und Mittel	Zur Überprüfung bedarf es (über die Beobachtung und Deutung hinaus) einer Deutungsbegründung
	Beurteilungstermini	Interessen i.e.S.: Wahrheit, Zufall	
	Reflexionstermini	Interessen i.w.S.: Werte und Einstellungen	

Tab. 3.4.2: Überblick über die Termini der psychologischen Wissenschaftssprache

Allgemein können wir festhalten, dass das methodisch kontrollierte Verstehen von Handlungen, Sprechhandlungen oder Texten in der systematischen Rekonstruktion ihrer relevanten Kontexte besteht und die Form einer Erzählung hat.

Systematische Rekonstruktion von Kontexten

Letzteres gilt bereits für die Klärung der potentiell kontextabhängigen Verwendungsweise der in der Erzählung vorkommenden Wörter – sowohl dafür, wie sie vom Sprecher gebraucht, als auch dafür, wie sie vom Angesprochenen verstanden werden. Um zu verstehen, dass und warum jemand (A), der vorher nach dem Wohlergehen seiner Frau befragt worden war, die Frage nach dem Wohlergehen seiner Familie jetzt nur auf seine Kinder be-

zieht, muss man sich eben diese Geschichte vergegenwärtigen (vgl. Tab. 3.4.3).

Tab. 3.4.3: Narrative Klärung kontextabhängiger Wortverwendungen

t ₁ = objektseitig	Herrn As Familie besteht aus seiner Frau, zwei Söhnen und einer Tochter
t ₂	Herr A hat die Frage nach dem Wohlergehen seiner Frau gerade schon beantwortet
t ₃ = subjektseitig	Herr A versteht die nachgeschobene Frage nach dem Wohlergehen seiner Familie jetzt nur noch als Frage nach dem Wohlergehen seiner Kinder

Auf die Frage, welches die relevanten Kontexte sind – d.h. welche von vielen möglichen Geschichten im konkreten Fall als erklärungskräftig angesehen werden können – ist eine schematische Antwort schon deshalb nicht möglich, weil Geschichten aus (mindestens) dreierlei Perspektiven erzählt werden können (vgl. Werbik & Appelsmeyer, 1999): aus der Perspektive des erlebenden Subjektes (1. Person), aus der Perspektive eines Interaktionspartners (2. Person) und aus der Perspektive des Beobachters (3. Person).

Eine zufriedenstellende psychologische Erklärung wird stets alle drei Perspektiven miteinander verbinden und gleichsam eine Metageschichte (3. Person) erzählen, welche die Geschichten 1. und 2. Person in einen systematisch-theoretischen Kontext stellt.

3.4.5 Metakommunikative Axiome

Funktionen der Sprechhandlung

Bereits Karl Bühler (1934) unterscheidet in seinem Hauptwerk *Sprachtheorie* drei Funktionen der Sprechhandlung:

- Darstellung (gegenstandsbezogene Perspektive),
- Ausdruck (senderbezogene Perspektive) und
- Appell (empfängerbezogene Perspektive).

Wenngleich sein Ansatz zunächst singulär blieb und erst Jahrzehnte später traditionsstiftend und schulbildend in den verschiedensten geistes- und sozialwissenschaftlichen Disziplinen wirkte, dominiert diese Bühler'sche Dreiteilung die Funktionszuweisungen sprachlicher Handlungen in den unterschiedlichsten theoretischen Ansätzen bis in die Gegenwart. In der psychologischen Kommunikationstheorie wurde er u.a. von Watzlawick, Beavin & Jackson (1967, 1969) und Schulz von Thun (1981) aufgegriffen.

Watzlawick et al. unterscheiden innerhalb der Kommunikation zwischen der einzelnen Mitteilung (Kommunikation i.e.S.) und dem wechselseitigen Austausch von Mitteilungen zwischen zwei oder mehreren Kommunikatoren, die sie als Interaktion bezeichnen. Darauf aufbauend, formulieren die Autoren fünf metakommunikative Axiome:

Die Axiome von Watzlawick et al.

1. Man kann nicht *nicht* kommunizieren
2. Jede Kommunikation hat einen Inhalts- und einen Beziehungsaspekt, und zwar derart, dass letzterer den ersten bestimmt und daher eine Metakommunikation ist.
3. Die Natur einer Beziehung ist durch die Interpunktion der Kommunikationsabläufe seitens der Partner bedingt.
4. Menschliche Kommunikation bedient sich digitaler und analoger Modalitäten. Digitale Kommunikationen haben eine komplexe und vielseitige logische Syntax, aber eine auf dem Gebiet der Beziehungen unzulängliche Semantik. Analoge Kommunikationen dagegen besitzen dieses semantische Potential, er-mangeln aber der für eindeutige Kommunikationen erforderlichen Syntax.
5. Zwischenmenschliche Kommunikationsabläufe sind entweder symmetrisch oder komplementär, je nachdem, ob die Beziehung zwischen den Partnern auf Gleichheit oder Unterschiedlichkeit beruht.

Mit dem ersten Axiom verweisen die Autoren darauf, dass Kommunikation nicht nur in verbalem Verhalten (bzw. Sprechhandlungen) besteht, sondern dass auch paraverbales und nonver-bales Verhalten Mitteilungscharakter haben. Da der Begriff des Verhaltens aber kein Gegenteil besitzt, kann man sich nicht *nicht* verhalten; auch

Axiom 1

"der Mann im überfüllten Wartesaal, der vor sich auf den Boden starrt oder mit geschlossenen Augen dasitzt, teilt den anderen mit, dass er weder sprechen noch angesprochen werden will, und gewöhnlich reagieren seine Nachbarn richtig darauf, indem sie ihn in Ruhe lassen. Dies ist nicht weniger ein Kommunikationsaustausch als ein angeregtes Gespräch" (Watzlawick et al., 1969, 51).

Wir fassen zusammen:

Man kann nicht nicht kommunizieren.

Mit dem zweiten Axiom greifen Watzlawick et al. die Bühler'sche Unterscheidung zwischen der Darstellungs- und der Ausdrucks-

Axiom 2

funktion von Sprechhandlungen auf, die sie als Inhalts- (report aspect) und Beziehungsaspekt (command aspect) der Kommunikation bezeichnen. Die Appellfunktion von Sprechhandlungen wird als Handlungsaspekt (parade aspect) ebenfalls unterschieden, aber im Folgenden nicht weiter analysiert. Schulz von Thun (1981) differenziert die Ausdrucksfunktion in einen Beziehungs- und einen Selbstmanifestationsaspekt.

Als Beispiel verweist Schulz von Thun auf einen Mann und eine Frau, die gemeinsam im Auto fahren. Der Mann sitzt am Steuer. Plötzlich schreit die Frau auf: "Die Ampel ist rot!".

- Unter dem Inhaltsaspekt betrachtet, handelt es sich um eine Mitteilung über die Ampelfarbe.
- Unter dem Selbstmanifestationsaspekt betrachtet, lautet die Mitteilung möglicherweise: "Ich habe schlechte Nerven."
- Unter dem Beziehungsaspekt mag sie als Kritik zu verstehen sein: "Dass ich immer für dich aufpassen muss!", und
- der Handlungsaspekt der Mitteilung besteht (u.a.) in dem Appell: "Tritt auf die Bremse!"

Um einem verbreiteten Missverständnis vorzubeugen, sei darauf hingewiesen, dass auch die Beziehung zwischen den Kommunikationspartnern, die Selbstmanifestation eines der Kommunikationspartner und/oder der an den anderen gerichtete Appell über die Darstellungsfunktion der Kommunikation ausgedrückt werden kann, indem die Sätze "Ich habe schlechte Nerven.", "Immer muss ich für dich aufpassen!" oder "Tritt auf die Bremse!" explizit ausgesprochen werden.

Die verschiedenen Aspekte derselben Mitteilung können durchaus widersprüchlich sein, z.B. wenn jemand mit Grabesmiene verkündet, wie sehr er sich freut.

Der zentrale Gedanke bereits bei Watzlawick et al. besteht darin,

- dass Kommunikationen bezüglich jeder der drei Bühler'schen Funktionen mehrdeutig sein können, und
- dass sich die verschiedenen Aspekte einer Kommunikation wechselseitig in ihrer Bedeutung bestimmen,
- also gleichzeitig einen metakommunikativen Gehalt haben, der bestimmt, wie die Kommunikation unter den jeweils anderen Aspekten zu verstehen ist.

Watzlawick et al. fokussieren im Weiteren zwar den metakommunikativen Gehalt des Beziehungsaspektes, weisen in einer Fußnote (1969, 56) aber darauf hin, dass dies arbiträr und lediglich von ihrem Forschungsinteresse geleitet ist. Von dieser Einschränkung befreit, können wir das zweite Axiom wie folgt zusammenfassen:

Jede Kommunikation erfüllt eine Darstellungsfunktion (Inhaltsaspekt), eine Ausdrucksfunktion (Selbstmanifestations- und Beziehungsaspekt) und eine Appellfunktion (Handlungsaspekt) derart, dass diese einander wechselseitig interpretieren und daher zugleich einen metakommunikativen Gehalt haben.

Unter Verweis auf die Kontextabhängigkeit von Bedeutungen können wir dies noch weiter präzisieren, indem wir festhalten:

Die Bedeutung einer Textstelle ergibt sich bezüglich jeder einzelnen dieser drei Funktionen aus ihrem Kontext, wobei zwischen ihrem externen Kontext (das sind (a) der Gegenstand, über den kommuniziert wird, und (b) die Situation, in der die Kommunikation stattfindet) und ihrem internen Kontext (das sind (a) der Text als Ganzes [interner Makro-Kontext] und (b) die jeweils anderen beiden Funktionen, unter denen die Textstelle zu verstehen ist [interner Mikro-Kontext]) zu unterscheiden ist.

Ein Beispiel: Zwei junge Frauen sitzen im Kaffee Sacher in Wien. Die eine schlank bis zur Grenze der Magersucht, die andere dick und behäbig. Sagt die eine "Wenn's das Sacher nicht gäb', wär' ich noch dünner". Sagt die andere "Ich auch". Schon das Wörtchen „noch' wird hier in zwei verschiedenen Bedeutungen gebraucht, und wie es gemeint ist, ergibt sich erst aus dem Kontext der Gewichtsprobleme der beiden (Gegenstand der Kommunikation).

Ein anderes Beispiel: Der Satz "Die Ampel ist rot!" wird in seiner Appellfunktion ganz unterschiedlich zu verstehen sein; je nachdem, ob es sich bei den beiden Personen im Auto um ein Liebespärrchen handelt, das während des Fahrens in heftige Knutscherei verwickelt ist, oder um ein genervtes Ehepaar, das gerade einen Streit hinter sich hat (Situation). Das eine Mal ist es nur der Appell, einen Unfall zu vermeiden, das andere Mal ist es – zumindest potentiell – auch der Appell, den Streit wieder aufzunehmen.

Axiom 3

Mit dem dritten Axiom verweisen Watzlawick et al. einerseits auf die Selbstreferentialität geschlossener Systeme (Kreisförmigkeit der Kommunikation), andererseits auf die systematische Perspektivendivergenz zwischen den Interaktionspartnern (vgl. divergierende Interpunktion), sowie drittens darauf, dass die Konstruktion der Rollen, welche die Interaktionspartner spielen, sozialen Regeln folgt, d.h. durch die Bereitwilligkeit der Teilnehmer entsteht, eine bestimmte Interpunktion anzunehmen:

"Die Versuchsratte, die sagte: „Ich habe meinen Versuchsleiter so abgerichtet, dass er jedes Mal, wenn ich den Hebel drücke, mir zu fressen gibt“ weigerte sich, die Interpunktion anzunehmen, die der Versuchsleiter ihr aufzuzwingen versuchte" (Whorf, 1956, zit.n. Watzlawick et al., 1969, 57f).

Diskrepanzen auf dem Gebiet der Interpunktion sind nach Watzlawick et al. die Wurzel vieler Beziehungskonflikte, wofür die Autoren ein in der Familientherapie nicht selten auftretendes Eheproblem als Beispiel anführen: Das Problem besteht darin, dass der Mann eine im wesentlichen passiv-zurückgezogene Haltung an den Tag legt, während seine Frau zu übertriebenem Nörgeln neigt. Im gemeinsamen Interview beschreibt der Mann seine Haltung als einzig mögliche Verteidigung gegen ihr Nörgeln

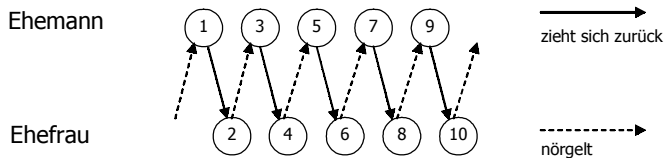
"Ich meide dich, weil du nörgelst",

während dies für sie eine krasse Entstellung dessen ist, was in ihrer Ehe "wirklich" vorgeht: dass nämlich der einzige Grund für ihre Kritik seine Absonderung von ihr ist:

"Ich nörgle, weil du mich meidest".

In dem endlos sich wiederholenden Hin und Her der gegenseitigen Vorwürfe und Selbstverteidigungen nimmt die Frau nur die Triaden 1-2-3, 3-4-5, 5-6-7 usw. wahr, in denen ihr Verhalten nur die Reaktion auf sein Verhalten ist, der Mann nimmt dagegen nur die Triaden 2-3-4, 4-5-6, 6-7-8 usw. wahr und sieht sein Verhalten nur als Reaktion auf das Verhalten der Frau (vgl. Abb. 3.4.3).

Abb. 3.4.3: Divergierende Interpunktionen eines Beziehungskonfliktes



Für die Analyse von Sprechhandlungen und Texten folgt daraus:

1. Kommunikationen sind selbstreferentiell.
2. Wegen der systematischen Perspektivendivergenz zwischen den Kommunikationspartnern ist der Kontext, in dem ein Text/eine Textstelle zu verstehen ist, nicht derselbe für Sprecher (1. Person) und Angesprochenen (2. Person). Während man eigene Handlungen (auch Sprechhandlungen) von der damit verbundenen Intention her versteht, erfährt man fremde Handlungen von ihren Wirkungen her. Hinzu kommen unterschiedliche Erfahrungshintergründe, Biographien, Interessen etc. Dies alles bewirkt, dass dieselbe Handlung/derselbe Text für Sender und Empfänger verschiedene, eventuell sogar entgegengesetzte subjektive Bedeutungen hat.
3. Texte haben sowohl einen manifesten als auch einen latenten Gehalt. Der manifeste Gehalt meint die (soziale) Wirklichkeit, die durch den Text dargestellt wird. Der latente Gehalt meint die (sozialen) Regeln, nach denen diese Wirklichkeit konstruiert wird und die den Kommunikationspartnern in der Regel nicht bewusst sind.

Mit dem vierten Axiom heben Watzlawick et al. auf einen wichtigen Unterschied zwischen verbaler Kommunikation einerseits und paraverbaler und nonverbaler Kommunikation andererseits ab. Axiom 4

Digitale Kommunikation (= verbale Kommunikation, idealiter in einer geklärten Wissenschaftssprache) hat eine vielseitige logische Syntax, die sie für die Darstellung des Inhaltsaspektes der Kommunikation prädestiniert.

Analoge Kommunikation (paraverbale und nonverbale Kommunikation) ermangelt dieser für eine eindeutige Kommunikation erforderlichen logischen Syntax, besitzt aber eine vielseitige Semantik, um dem Selbstmanifestations- und Beziehungsaspekt Ausdruck zu verleihen und/oder die Appellfunktion der Kommunikation auszuüben.

Daraus ergibt sich als wichtige Konsequenz, dass die Darstellung des Inhaltsaspektes eines Textes in der Regel eindeutiger und weniger interpretationsoffen ist als der Ausdruck seines Selbstmanifestations- und Beziehungsaspektes und/oder der von ihm ausgeübte Appell.

Axiom 5

Das fünfte Axiom ist vor allem für die Analyse der Appellfunktion von Sprechhandlungen und Texten bedeutsam und verweist darauf, dass der von einem Text ausgehende Appell von der Art der Beziehung zwischen den Kommunikationspartnern abhängt, die entweder symmetrisch oder komplementär sein kann.

- Im ersten Fall ist das Verhalten der Kommunikationspartner sozusagen spiegelbildlich. Symmetrische Interaktionsabläufe zeichnen sich durch Streben nach Gleichheit und Verminderung von Unterschieden aus: Gleichem wird Gleiches zur Seite gestellt.
- Komplementäre Interaktionen sind dagegen durch einander ergänzende Unterschiedlichkeiten gekennzeichnet. Komplementäre Beziehungen beruhen auf gesellschaftlichen und kulturellen Kontexten (wie z.B. Mutter und Kind, Arzt und Patient, Lehrer und Schüler), wobei es nicht etwa so ist, dass ein Partner dem anderen eine komplementäre Beziehung aufzwingt, sondern *beide* verhalten sich in einer Weise, die das jeweilige Verhalten des anderen voraussetzt, es gleichzeitig aber auch bedingt.

Damit sind zwei Grundformen sozialer Beziehungen angesprochen, für welche unterschiedliche Regelsysteme gelten.

Für die Textanalyse ergibt sich daraus:

Symmetrie und Komplementarität stellen zwei Grundformen sozialer Beziehungen dar, die ihren je eigenen, sozial und kulturell vermittelten Regeln folgen, die bei der Analyse der Appellfunktion von Texten in Rechnung zu stellen sind.

Kontrollfragen

- Was versteht man unter dem Ergebnis einer Handlung?
- Worin besteht der Unterschied zwischen Zwecken und Handlungsregeln?
- Erläutern Sie, warum auch Handlungsregeln finale Handlungsorientierungen sind!
- Worin besteht der Unterschied zwischen finalen und nonfinalen Anforderungen bzw. Handlungsorientierungen?
- Welche Arten von nonfinalen Handlungsorientierungen lassen sich unterscheiden?
- Wie können Interessen i.e.S. definiert werden, und inwiefern sind sie zwar nonfinal, aber doch quasi-final?
- Erläutern Sie anhand eines Beispiels, warum Interessen i.w.S. nicht

durch Konstruktionsprinzipien definiert werden können!

- Erläutern Sie, warum Interessen i.w.S. nicht quasi-final sind!
- Erläutern Sie, worin sich die Interessen i.w.S. von Handlungszielen und Interessen i.e.S. unterscheiden!
- Zeigen Sie am Beispiel der Gewaltfreiheit, dass es sich bei Lebensorientierungen
 - a) um afinale Orientierungen handelt, die
 - b) ein an den Rändern offenes Bedeutungsfeld besitzen!
- Erläutern Sie den Unterschied zwischen personaler und struktureller Gewalt!
- Welcher Begründungszusammenhang besteht bei Galtung zwischen der Erweiterung des Gewaltbegriffs einerseits und der Formulierung eines positiven Friedensbegriffes andererseits?
- Diskutieren Sie die gewaltfördernde Valenz des Skandalisierungs- und Diskreditierungspotential des Gewaltbegriffs!
- Worin besteht die perspektivische Verwendungsweise des Gewaltbegriffs?
- Worin besteht das Reziprozitätsprinzip?
- Begründen Sie, warum die Mischung aus Perspektivendivergenz und Reziprozitätsprinzip die Gewalteskalation fördert!
- Erläutern Sie anhand eines Beispiels, warum und inwiefern die Forderung nach einer eindeutig geklärten und kontextunabhängigen Terminologie im Falle der Lebensorientierungen relativiert werden muss!
- Erläutern Sie anhand eines Beispiels
 - a) die Kontextabhängigkeit von Bedeutungen sowie
 - b) die Rekonstruktion von Bedeutungen nach dem Schema der narrativen Erklärung!
- Geben Sie einen Überblick über die Termini der psychologischen Wissenschaftssprache und erläutern Sie, was erforderlich ist, um mit ihnen gebildete Aussagen zu überprüfen!
- Erläutern Sie, warum das methodisch kontrollierte Verstehen ein zentrales Methodenproblem der Psychologie darstellt!
- Erläutern Sie anhand von Beispielen, worin das methodisch kontrollierte Verstehen von Handlungen, Sprechhandlungen und/oder Texten besteht und welche Form es hat!
- Welche Funktionen der Sprechhandlung werden von Bühler unterschieden?
- Wie lauten die metakommunikativen Axiome von Watzlawick et al.?
- Welche Aspekte der Kommunikation werden von Watzlawick et al. unterschieden? In welcher Beziehung stehen sie zu den Funktionen der Sprechhandlung nach Bühler, und welche Differenzierung dieser Aspekte wurde durch Schulz von Thun vorgenommen?
- Erläutern Sie die verschiedenen Aspekte der Kommunikation anhand eines Beispiels!
- Diskutieren Sie, ob bzw. in welcher Weise die Beziehung zwischen den Kommunikationspartnern, die Selbstmanifestation des Sprechers und/oder der an den Angesprochenen gerichtete Appell über die Darstellungsfunktion der Kommunikation ausgedrückt werden kann!
- Erläutern Sie, worin der metakommunikative Gehalt der verschiedenen Funktionen der Sprechhandlung besteht!

- Erläutern Sie, woraus sich die Bedeutung einer Textstelle ergibt! Geben Sie Beispiele!
- Erläutern Sie anhand eines Beispiels die Termini
 - a) ‚Selbstreferentialität‘ geschlossener Systeme,
 - b) ‚Perspektivendivergenz‘ der Kommunikationspartner und
 - c) ‚Interpunktion‘ der Kommunikation!
- Worin besteht der Unterschied zwischen dem manifesten und dem latenten Gehalt eines Textes?
- Worin besteht der Unterschied zwischen digitaler und analoger Kommunikation? Welche Konsequenzen ergeben sich daraus für die Eindeutigkeit des Inhalts-, des Selbstmanifestations-, des Beziehungs- und des Handlungsaspektes der Kommunikation?
- Worin besteht der Unterschied zwischen symmetrischen und komplementären Kommunikationen?

3.5

Die systemtheoretische Erweiterung des handlungstheoretischen Modells

3.5.1 Der Begriff des Systems

Mit der kommunikationstheoretischen Betrachtung zwischenmenschlicher Interaktionen hat das handlungstheoretische Modell insofern eine wesentliche Erweiterung erfahren, als wir nicht mehr das Handeln isolierter Subjekte betrachten, sondern die Beziehung zwischen den Interaktionspartnern als ein System darstellen, das seinen eigenen Gesetzmäßigkeiten folgt.

Das System

So erzählt im Beispiel des von Watzlawick et al. beschriebenen Eheproblems der Mann, wie die Frau an ihm herumnörgelt (handlungsauslösende Konstellation), dass er den ständigen Streit vermeiden will (Intention), dass er meint, im Hobbykeller seine Ruhe finden zu können (Miteinschätzung) und also den Großteil seiner Freizeit allein im Hobbykeller verbringt (Verhalten) (vgl. Abb. 3.5.1).

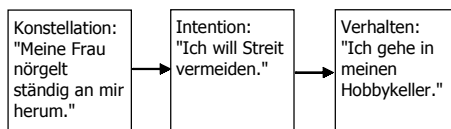


Abb. 3.5.1: Die subjektive Wirklichkeit des Mannes

Die Frau erzählt, wie sich der Mann immer zurückzieht (handlungsauslösende Konstellation), dass es ihr nur darum geht, die Beziehung wieder herzustellen (Intention), dass sie meint mit ihm darüber reden zu müssen (Miteinschätzung) und wie sie also versucht, ein Gespräch darüber anzufangen, dass es so nicht weitergehen kann (Verhalten) (vgl. Abb. 3.5.2).

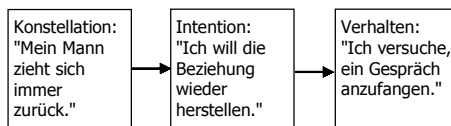


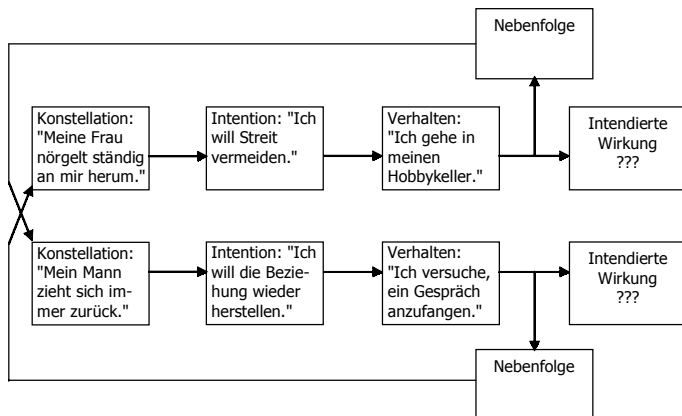
Abb. 3.5.2: Die subjektive Wirklichkeit der Frau

Der autonome Prozess

Der Familientherapeut sieht, wie die Interaktion zwischen den Ehepartnern einen autonomen Prozess (Meier, 1978) darstellt, dessen Eigendynamik vom Handlungserfolg der Beteiligten unabhängig ist und allein durch die (nicht intendierten) Nebenfolgen ihres Handelns angetrieben wird (Hoyningen-Huene, 1983). Unabhängig davon, ob es dem Mann gelingt, in der aktuellen Situation Streit zu vermeiden oder nicht; er hat den Streit bestenfalls aufgeschoben. Und unabhängig davon, ob es der Frau gelingt, die Kommunikation mit ihrem Mann für den Moment wieder herzustellen oder nicht; sie hat seinen Rückzug damit bestenfalls hinausgezögert (vgl. Abb. 3.5.3).

Während jeder von ihnen den Konflikt so interpunktiert, dass er nur auf das Verhalten des anderen reagiert, reproduzieren die Nebenfolgen ihres Verhaltens genau jene Konstellation, die gerade das Verhalten des Partners auslöst, das sie vermeiden möchten und gegen das sie sich verteidigen zu müssen glauben.

Abb. 3.5.3: Die objektive Wirklichkeit des Systems



3.5.2 Selbstreferentialität

Drei Wirklichkeiten

Mit der Analyse der Interaktion zwischen den Ehepartnern hat sich in methodologischer Hinsicht einiges geändert. Wir haben es nicht mehr nur mit einer Wirklichkeit zu tun, sondern gleich mit drei Wirklichkeiten. Zunächst:

- mit der subjektseitig definierten Wirklichkeit, welche der Konflikt für den Ehemann hat, und
- mit der subjektseitig definierten Wirklichkeit, welche der Kon-

flikt für die Ehefrau hat.

Diese beiden Wirklichkeiten resultieren aus der je eigenen Verstrickung in den Konflikt und können nur vom Innenstandpunkt der Konfliktparteien selbst her richtig erfasst werden.

Drittens hat der Konflikt aber auch eine objektseitig definierte Wirklichkeit, die nur von einem Außenstandpunkt aus erfassbar ist und die beschreibt, wie die subjektseitig definierten Wirklichkeiten miteinander interagieren. Diese objektseitig definierte Wirklichkeit betrifft nicht mehr das Handeln der einen oder anderen Konfliktpartei für sich, sondern das Verhalten des von ihnen gebildeten Systems. Solche Systeme sind jedoch grundsätzlich selbstreferentiell, d.h. das Verhalten des Systems zeitigt nicht nur irgendwelche Ergebnisse und/oder Wirkungen, welche außerhalb des Systems liegen und die Umweltgegebenheiten verändern, innerhalb derer das System operiert, sondern es findet eine Rückkopplung auf das System selbst statt.

Dies hat weitreichende methodologische Konsequenzen, die sich mit Hilfe des auf den englischen Mathematiker und Logiker Alan Turing (1936/37) zurückgehenden Begriffs der Maschine erläutern lassen. Triviale und nichttriviale Maschinen

Der Terminus ‚Maschine‘ meint dabei nicht ein physisches Arrangement von Knöpfen, Hebeln, Zahnrädern, Mikrochips etc., sondern eine Liste von Regeln, wobei es unwesentlich ist, ob es sich um empirische oder um strukturelle Regeln handelt. Man unterscheidet zwei Arten solcher Maschinen: triviale und nichttriviale Maschinen.

Triviale Maschinen sind dadurch definiert, dass sie *genau eine*, und zwar unveränderliche Regel haben. Als Beispiel für eine triviale Maschine können wir etwa an das Weber'sche Gesetz denken. Die Regel lautet hier:

$$\begin{aligned} |V-S| \geq c \times S &\rightarrow w(V-S) \\ |V-S| < c \times S &\rightarrow \neg w(V-S) \end{aligned}$$

Ein anderes Beispiel für triviale Maschinen sind die von Siegler beschriebenen Entwicklungsstufen bei der Lösung des Balkenwaagen-Problems. Dabei definiert jede der vier Stufen eine andere triviale Maschine, und je nachdem, welchen Aufgabentyp wir als Input (x) verwenden, produziert die jeweilige Maschine einen ganz bestimmten Output (y). Das Verhalten der Maschine ist analytisch bestimmbar und voraussagbar (vgl. Abb. 3.5.4).

Abb. 3.5.4: Triviale Maschine

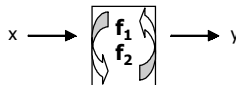


- a) synthetisch determiniert
- b) vergangenheitsunabhängig
- c) analytisch bestimmbar
- d) voraussagbar

Nichttriviale Maschinen sind dagegen dadurch gekennzeichnet, dass sie über *mehrere* Regeln verfügen. Sie kennen also mehrere innere Zustände, unter denen jeweils eine andere Regel operiert, und sie verfügen außerdem über eine Regel zweiter Ordnung, nach welcher sich der innere Zustand von einem Schritt zum nächsten verändert (vgl. Abb. 3.5.5).

Z.B. betrachten wir Maschinen zur Übersetzung der Symbole A, B, C, D in eine andere Reihenfolge derselben 4 Symbole. Dann gibt es $4! = 24$ verschiedene triviale Maschinen, die das leisten, bzw. 24 denkbare innere Zustände, die eine nichttriviale Maschine haben könnte. Die einfachste dieser nichttrivialen Maschinen ist eine, die nur zwei dieser inneren Zustände kennt und jedes Mal, wenn sie ein Symbol ersetzt hat, in den jeweils anderen Zustand wechselt (Regel 2. Ordnung).

Abb. 3.5.5: Nichttriviale Maschine



- a) synthetisch determiniert
- b) vergangenheitsabhängig
- c) analytisch unbestimmbar
- d) unvoraussagbar

Unsere Maschine kennt z.B. die inneren Zustände

f₁:	A → B	f₂:	A → C
	B → C		B → D
	C → D		C → A
	D → A		D → B

Befindet sich also die Maschine z.B. im Zustand f_1 und bietet man ihr den Input A, so produziert sie den Output B und wechselt in den Zustand f_2 . Bietet man ihr jetzt wieder den Input A, so produziert sie – anders als triviale Maschinen – jetzt aber nicht wieder den Output B, sondern stattdessen den Output C. Dadurch würde z.B. die Folge A, B, C, D je nachdem, in welchem Anfangszustand sich die Maschine befunden hat, als man sie zu untersuchen begann, das eine Mal in die Folge B, D, D, B (Anfangszustand = f_1) und das andere Mal in die Folge C, C, A, A (Anfangszustand = f_2) übersetzt.

Um das Verhalten der Maschine zu analysieren, müsste man die richtige Maschine aus

$$n_z(x,y) = y^{xz}$$

möglichen Maschinen herausfinden, die genau x Inputsymbole, bis zu y Outputsymbole und z innere Zustände kennt. In unserem Beispiel ist $x = y = 4$ und $z = 2$, also wäre die Maschine aus $n_2(4,4) = 4^{4 \times 2} = 65536$ möglichen Maschinen herauszufinden. Wir hätten also das Verhalten von 65536 Maschinen zu analysieren, um die richtige Maschine zu identifizieren. Dabei ist diese Anzahl noch relativ klein, weil bekannt war, dass die Maschine nur 2 innere Zustände kennt. Ist dies nicht bekannt, so muss man mit allen 24 möglichen inneren Zuständen rechnen und die Anzahl der Maschinen, deren Verhalten es zu analysieren gilt, erhöht sich auf $4^{4 \times 24} = 2^{192}$. Diese Zahl bewegt sich in der Größenordnung von 10^{58} und ist damit größer als das Alter unseres Universums in Mikrosekunden, das gerade mal $10^{27} \mu\text{s}$ beträgt. Von Foerster (1994) schließt daraus, dass bereits das Verhalten so einfacher Klassen von nichttrivialen Maschinen nicht mehr berechenbar, d.h. transcomputational ist und verweist auf Arthur Gill (1962), der in seiner *Introduction to the Theory of Finite State Machines* gezeigt hat, dass es keineswegs einer komplizierten inneren Organisation solcher Maschinen bedarf, um es unmöglich zu machen, die Maschine in einer endlichen Reihe von Experimenten zu identifizieren.

Die Systemtheorie geht noch einen Schritt weiter, indem sie operational geschlossene Systeme untersucht, bei denen der Output (y) zum nächsten Input x' wird. In diesem Sinne werden z.B. beim autonomen Prozess die Nebenfolgen (Output) zum neuen Input (Reproduktion der handlungsauslösenden Konstellation).

Geschlossene Systeme

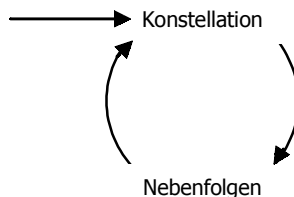


Abb. 3.5.6: Kreisförmigkeit des autonomen Prozesses

Eine wichtige Eigenschaft solcher geschlossener Systeme ist, dass sie unter bestimmten Bedingungen zu einem stabilen Eigen-

verhalten konvergieren. Im Falle des autonomen Prozesses ist das trivial: weil das System hier nur einen inneren Zustand (Konkurrenz) und – darüber hinaus – nur einen Output kennt, der zum neuen Input wird. Das System verfällt in einen endlosen Kreislauf, sobald die handlungsauslösende Konstellation ein erstes Mal als Input aufgetreten ist.

Etwas komplexer ist das Verhalten der oben beschriebenen nicht-trivialen Maschine, die bei operativer Schließung unabhängig davon, in welchem Anfangszustand sie sich befindet und welches Symbol man als Input verwendet, in die sich stets wiederholende Zeichenabfolge A, B, D, A, C, D, B, C verfällt.

Solche einfachen Beispiele haben – wie v. Foerster zu Recht bemerkt – natürlich nur metaphorischen Charakter. Schon bei der Untersuchung zwischenmenschlicher Interaktionen müssen wir von (mindestens) zwei Klassen von handlungsauslösenden Konstellationen (Inputs) ausgehen (konfliktartig oder konfliktfrei) und (mindestens) zwei innere Zustände des Systems (Konkurrenz oder Kooperation) unterscheiden. Zum anderen verfügt das System über mehrere mögliche Outputs (Handlungswirkungen und Nebenfolgen), über die – nach unserer bisherigen Analyse – nur ausgesagt ist, dass der Zustand der Konkurrenz Konfliktkonstellationen (als neuen Input) reproduziert. Drittens muss davon ausgegangen werden, dass der Output auch eine Rückkoppelung auf den inneren Zustand hat, in dem sich das System im nächsten Schritt befindet, usw.

Dennoch sind solche einfachen Beispiele geeignet, um zwei Grundprinzipien zu veranschaulichen, die bei der Untersuchung komplexer Systeme in Rechnung zu stellen sind:

- Erstens, dass es komplexe Systeme durch ihr Eigenverhalten zwar gestatten, zu beobachten, *was* sie tun, es aber wegen ihrer analytischen Unerreichbarkeit unmöglich machen, zu wissen, *wie* sie es tun (v. Foerster, 1994). Zwar lassen sich ggf. Kalküle konstruieren, die das Verhalten des Systems simulieren; es lässt sich aber nicht mit Sicherheit entscheiden, ob das System tatsächlich ein Modell dieses Kalküls ist oder nach einem ganz anderen Kalkül funktioniert, der dasselbe Verhalten zeigt.
- Das zweite Grundprinzip lautet, dass wir es in der Systemtheorie grundsätzlich mit in der Zeit ablaufenden Prozessen zu tun haben. Entsprechend hat bereits der Begründer der Kyberne-

mindestens eine Wenn-dann-Beziehung besteht. Die Pfeile heißen Wirkungslinien und stellen das Zusammenwirken der Blöcke dar. Die Pfeilspitze gibt die Wirkungsrichtung an. Wirkungslinien, die die Hüllkurve von außen nach innen durchdringen, symbolisieren freie Eingänge des Systems. Durchdringen sie die Hüllkurve von innen nach außen, so symbolisieren sie offene Ausgänge.

Die Analyse von Systemen ist grundsätzlich *rekursiv*, d.h. die einzelnen Blöcke können ihrerseits wiederum als Systeme betrachtet werden. Der freie Ausgang des Systems (x), der in Übereinstimmung mit dem Sollwert gehalten werden soll, heißt Regelgröße. Der Block, der diese Größe erzeugt, wird als Regelstrecke bezeichnet. Der Block, der die Regelstrecke in Abhängigkeit von der Soll-Ist-Differenz aussteuert, heißt Regler. Die Variable (u), über die das geschieht, wird Stellgröße genannt. Die freien Eingänge (v), die die Regelgröße ebenfalls beeinflussen, heißen Störgrößen.

Stabile und instabile Regelkreise

Wir können dieses Beispiel jetzt benutzen, um das Zeitverhalten von Regelkreisen zu studieren. Dabei ist zu beachten, dass die Werte seiner Variablen zu jedem Zeitpunkt angebar sind und sich zudem für jeden einzelnen Block eine Funktion angeben lässt. In unserem Beispiel: $u = f(z)$, $x = g(u,v)$ und $y = h(x)$. Sofern der abgebildete Regelkreis stabil ist und die letzten Änderungen der freien Eingänge hinreichend lange zurückliegen, befindet sich das System in einem eingeschwungenen Zustand. D.h., jede Variable hat einen festen Wert, der sich in der Zeit nicht mehr ändert. Öffnen wir aber nun z.B. für 5 Minuten das Fenster, so tritt eine Änderung der Störgrößen ($v \rightarrow v'$) ein: Es kommt kalte Luft ins Zimmer. Die Regelstrecke (Zimmer mit Heizkörper) verändert den Wert der Regelgröße (Temperatur) $x \rightarrow x' = g(u,v')$. Der Messfühler errechnet den veränderten Ist-Wert $y \rightarrow y' = h(x')$, und der Vergleicher stellt eine Soll-Ist-Differenz $z \rightarrow z' = y' - w < 0$ fest (es ist kälter als es sein sollte). Der Regler erhöht die Warmwasserzufuhr in den Heizkörper $u \rightarrow u' = f(z')$, wodurch sich wiederum die Regelstrecke verändert $x' \rightarrow x'' = g(u',v')$, d.h. es wird wärmer im Zimmer usw.

Dabei ist es aber keineswegs selbstverständlich, dass sich der Regelkreis wieder in einen stabilen Zustand einschwingt. Es wäre auch der Fall denkbar, dass sich der Heizkörper so überhitzt, dass er noch weiter heizt, wenn die Solltemperatur längst erreicht ist und der Regler den Wasserzufluss schon lange geschlossen hat. Das Zimmer wird immer wärmer, die Regelung schwingt über.

Das Zimmer kühlt dann langsam wieder ab, bis die Solltemperatur wieder unterschritten wird und der Regler den Warmwasserzufluss wieder öffnet. Es dauert jedoch seine Zeit, bis der Heizkörper wieder warm wird. Die Zimmertemperatur sinkt also zunächst noch weiter ab, um schließlich in derselben Weise wie vorher wieder über den Sollwert hinauszuschließen. Das kann sich theoretisch unendlich wiederholen, und der Regelkreis kann also Eigenschaften haben, bei denen er nur noch schwingt und nicht mehr zur Ruhe kommt.

Wie man zeigen kann, sind es zwei Bedingungen, die zu derart instabilen Regelkreisen führen: erstens die Trägheit der Systemkomponenten (in unserem Beispiel, dass sich der Heizkörper zu langsam erhitzt bzw. abkühlt), und zweitens der ‚Verstärkungsfaktor‘, d.h., dass ein Regelkreis umso mehr zum Schwingen neigt, je stärker der Regler der Störung entgegenwirkt.

Glaser (2000, 77) verweist in diesem Zusammenhang darauf, dass dieser Satz dem naiven Alltagsdenken widerspricht, wonach "ein Regelkreis umso ‚leistungsfähiger‘ wird, je stärker die Rückkoppelung der Störung entgegenarbeitet. Vor allem die Simulationsversuche Dietrich Dörners zum Problemlösen in komplexen Systemen zeigen dies. Die erfolglosen Versuchspersonen greifen in Regelkreise zu massiv, also mit hohem Verstärkungsfaktor ein, weil sie glauben, man müsse deutlichen Missständen mit äußerst wirksamen Maßnahmen begegnen (z.B. Dörner 1988, S. 200ff). Dabei machen sie die simulierten Wirtschafts- und Ökosysteme entweder instabil oder bringen sie schnell in nicht mehr umkehrbare, katastrophale Grenzlagen, die sich als Überschwngen bis zur Systemzerstörung interpretieren lassen."

3.5.4 Schärfegrade des autonomen Prozesses

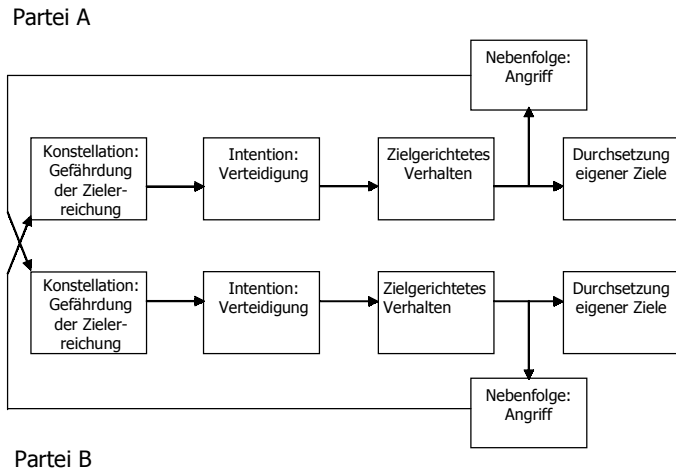
Das Konzept des autonomen Prozesses wurde von dem Historiker Christian Meier (1978) zur Erklärung des Unterganges des römischen Reiches entwickelt und von Hoyningen-Huene (1983) in ein kybernetisches Modell übersetzt. Die Erweiterung des Konzeptes zur Analyse von Konflikten geht zurück auf Kempf (1993).

Die autonome Dynamik kompetitiver Prozesse

Dabei zeigt sich, dass die autonome Konfliktdynamik von den konkreten Inhalten des Konfliktes unabhängig ist und sich immer dann einstellt, wenn ein Konflikt als kompetitiver Prozess nach dem win-lose-Prinzip ausgeglichen wird (Kempf, 2000). Da win-lose bedeu-

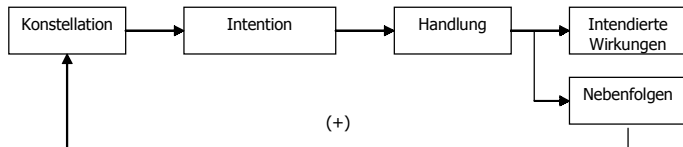
tet, dass immer nur einer gewinnen kann und der andere verlieren muss, hat in einer solchen Situation jede Durchsetzung eigener Ziele die Nebenfolge, dass damit die Zielerreichung des anderen verhindert wird, und jeder Versuch, sich gegen diesen Angriff zu verteidigen und die eigenen Ziele doch noch durchzusetzen, hat die Nebenfolge, ein (Gegen-)Angriff zu sein (vgl. Abb. 3.5.8).

Abb. 3.5.8: Die autonome Dynamik kompetitiver Prozesse



In seiner allgemeinsten Form kann der autonome Prozess durch das Regelkreismodell in Abb. 3.5.9 dargestellt werden. Die positive Rückkoppelung von den Nebenfolgen auf die handlungsauslösende Konstellation kann (wie im Falle des kompetitiven Konfliktes, vgl. Abb. 3.5.8) über einen anderen Akteur (oder über mehrere) vermittelt sein. Für den autonomen Prozess als solchen ist dies nicht entscheidend. Worauf es ankommt, ist lediglich, dass eine solche Rückkoppelung existiert und dass sie vom Handlungserfolg unabhängig ist.

Abb. 3.5.9: Der autonome Prozess



Hinsichtlich der Frage, wie ein solcher autonomer Prozess aufgelöst werden kann, unterscheidet Hoyningen-Huene drei Schärfegrade des autonomen Prozesses:

Der erste Schärfegrad ist gegeben, wenn die Auflösung durch einfachen Mittelaustausch erfolgen kann, d.h. wenn es eine andere Handlungsmöglichkeit gibt, die ebenfalls die intendierten Wirkungen hat, nicht jedoch die schädlichen Nebenfolgen (vgl. Abb. 3.5.10). In diesem Fall ist der autonome Prozess relativ leicht aufzulösen: Mit welchen Mitteln man seine Ziele erreicht, kann einem ziemlich egal sein.

Der erste Schärfegrad

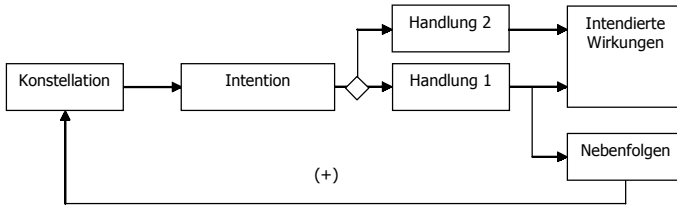


Abb. 3.5.10: Der erste Schärfegrad des autonomen Prozesses

Der zweite Schärfegrad ist erreicht, wenn ein Mittelaustausch nicht hilft, aber eine Intensionsänderung möglich ist (vgl. Abb. 3.5.11). Das ist schon weit schwieriger, denn dazu muss man etwas aufgeben, was man eigentlich wollte.

Der zweite Schärfegrad

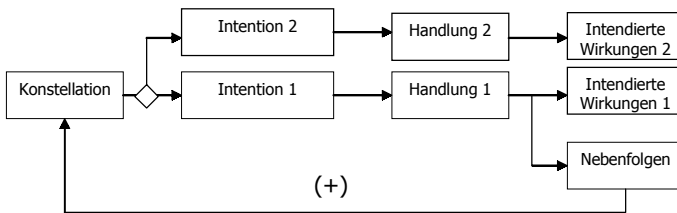


Abb. 3.5.11: Der zweite Schärfegrad des autonomen Prozesses

Beim dritten Schärfegrad ist auch dies nicht mehr möglich. Eine Auflösung des autonomen Prozesses ist nur noch möglich, wenn es gelingt, die Konstellation aufzubrechen, wofür grundsätzlich zwei Möglichkeiten bestehen:

Der dritte Schärfegrad

- Änderung der (objektseitig definierten) Umweltbedingungen (z.B. Beseitigung struktureller Gewalt), oder
- Änderung der subjektiven Interpretation dieser Umweltbedingungen, womit sich (subjektseitig) eine neue Situation ergibt, in der ggf. andere Situationsaspekte handlungsrelevant werden.

Diese Neuinterpretation der Umweltbedingungen ist am schwierigsten zu bewerkstelligen, weil sie an die Lebensorientierungen

geht und daher eine Identitätskrise einschließt: Orientierungen, die bisher selbstverständlich waren, müssen plötzlich zur Disposition gestellt werden.

Kontrollfragen

- Diskutieren Sie anhand eines Beispiels, inwiefern die systemtheoretische Betrachtungsweise eine Erweiterung des handlungstheoretischen Modells darstellt!
- Erläutern Sie anhand eines Beispiels den Begriff des autonomen Prozesses!
- Welche Arten von Wirklichkeit müssen bei der Analyse von Interaktionen betrachtet werden?
- Was bedeutet subjekt- und objektseitige Wirklichkeit?
- Was versteht man unter der Selbstreferentialität von Systemen?
- Was ist eine (Turing-)Maschine?
- Worin besteht der Unterschied zwischen trivialen und nichttrivialen Maschinen? Geben Sie Beispiele!
- Was bedeutet der Terminus ‚transcomputational‘?
- Was versteht man unter einem System?
- Was versteht man unter der operationalen Schließung von Systemen?
- Welche Konsequenzen ergeben sich aus ihrer Transcomputationalität für die Untersuchung komplexer Systeme?
- Welches sind die Elemente eines Blockdiagramms? Was stellen sie dar?
- Was versteht man darunter, dass sich ein System im eingeschwungenen Zustand befindet?
- Worin besteht der Unterschied zwischen stabilen und instabilen Systemen?
- Welches sind die Bedingungen, die zu instabilen Regelkreisen führen?
- Erläutern Sie, warum kompetitiv ausgetragenen Konflikte stets autonome Prozesse sind!
- Erläutern Sie,
 - a) welche Schärfegrade des autonomen Prozesses man unterscheiden kann,
 - b) welche Möglichkeiten es zum Aufbrechen des autonomen Prozesses gibt, und
 - c) warum dies mit zunehmendem Schärfegrad des Prozesses immer schwieriger wird!

3.6

Informationsverarbeitungsmodelle

3.6.1 Der Begriff der intervenierenden Variablen

Hatte der Behaviorismus versucht, menschliches Verhalten durch triviale Maschinen zu modellieren, welche eine direkte Übersetzung von Stimuli (S) in Reaktionen (R) beschreiben, so wurde das S-R-Modell im Neobehaviorismus zum S-O-R-Modell erweitert, das den Organismus (O) als vermittelnde Instanz zwischen Stimulus und Reaktion schaltet.

Der Organismus selbst wurde vom Neobehaviorismus als *black box* behandelt, und die Prozesse, die innerhalb dieser black box ablaufen, sollten durch intervenierende Variablen erklärt werden, die zwar selbst nicht beobachtbar, aber durch beobachtbare Referenzoperationen operationalisierbar sind.

Eine frühe Form dieser Art von Theorienbildung stellt die kognitive Lerntheorie von Edward C. Tolman (1937) dar, deren Grundannahmen wir wie folgt zusammenfassen können:

Die kognitive Lerntheorie von Edward C. Tolman

1. Lernprozesse entstehen als Ergebnis spezifischer räumlicher oder zeitlicher Anordnungen von Reizobjekten, die dem Organismus einmal oder wiederholt dargeboten werden.
2. Aus dieser Begegnung mit einer gegebenen Reizsequenz resultiert im Organismus ein bestimmter Zustand oder Prozess, der als Zeichen-Gestalt-Erwartung (sign-gestalt-expectation) bezeichnet wird.
3. Darunter wird eine Einstellung (set) oder ein Prozess im Organismus verstanden, der auf die Tatsache abgestimmt (adjusted) ist, dass ein räumlich oder zeitlich unmittelbar gegebenes Reizobjekt als Folge einer bestimmten Art von Verhaltensweisen des Organismus (oder ihres Ausbleibens) zu einem so und so beschaffenen räumlich und zeitlich entfernten Zielobjekt führen wird.
4. Derartige Einstellungen äußern sich dadurch, dass sich der Organismus unter geeigneten Motivationsbedingungen, die aus

dem Wert (value) des Zielobjektes resultieren, dem unmittelbaren Reizobjekt gegenüber so verhält, dass er dadurch entweder zum entfernteren Reizobjekt gelangt oder dieses vermeidet.

5. Wenn sich die Umweltbedingungen in der Weise verändern, dass das eine Reizobjekt samt dem gegebenen Verhalten nicht mehr zum anderen Reizobjekt führt, dann wird diese Einstellung mit dem resultierenden Verhalten unter normalen Umständen früher oder später wieder verschwinden (d.h. verlernt).
6. Die Gesetze und Prinzipien des Lernens sind Aussagen über jene Arten von Ereignissequenzen in der Umwelt und über die Bedingungen ihrer Darbietung, welche die Ausbildung korrespondierender Einstellungen (Zeichen-Gestalt-Erwartungen) begünstigen.

Ungeachtet ihrer intentional anmutenden Lautgestalt (value, expectation etc.) handelt es sich bei den intervenierenden Variablen jedoch *nicht* um Sinngehalte – die zwar ebenfalls Beobachtungskorrelate besitzen, aber nicht mit Beobachtbarem gleichgesetzt werden können –, sondern sie werden durch die entsprechenden Referenzoperationen (pointer readings) *operational definiert*. So wird z.B. der positive Wert eines Zielobjektes in Tolmans Theorie durch die Stärke der Tendenz bestimmt, mit der ein Organismus zu dieser Art von Zielobjekten strebt (vgl. Foppa, 1965, 360).

Der Verdacht der
Realitätsverdoppelung

Der Verdacht, dass es sich dabei um eine Spielart der Realitätsverdoppelung handeln könnte, ist nicht von der Hand zu weisen:

- Die Zeichen-Gestalt-Erwartung äußert sich im Verhalten des Organismus, durch welches er zu positiv wertigen Zielobjekten gelangt und/oder negativ wertige Zielobjekte meidet.
- Der Wert des Zielobjektes wird durch die Tendenz bestimmt, mit welcher der Organismus zu dieser Art von Zielobjekten hinstrebt (positiver Wert) oder von ihnen wegstrebt (negativer Wert).
- Diese Tendenz misst sich am Verhalten des Organismus, durch welches er entweder zu dieser Art von Zielobjekten gelangt (Appetenzverhalten) oder und durch welches er sie meidet (Aversionsverhalten).

Fazit: Der Organismus verhält sich so, dass er zu jenen Zielobjekten gelangt, zu denen er gelangt, und dass er jene Zielobjekte meidet, die er meidet. Zeichen-Gestalt-Erwartung, Wert des Ziel-

objektes und Appetenz- vs. Aversionsverhalten des Organismus sind inhaltsgleiche Beschreibungen desselben Sachverhaltes ohne jeden Erklärungsgehalt.

Ähnlich wie bei der Triebtheorie der Aggression (vgl. Kapitel 2.5.2.3) wird sich diese Realitätsverdoppelung nicht beheben lassen; es sei denn, dass man auf eine unmittelbare Operationalisierung der intervenierenden Variablen verzichtet und stattdessen hinreichend genaue Vorstellungen über die in der black box ablaufenden Prozesse entwickelt, so dass sich konkrete Prognosen über das Verhalten des Systems ableiten lassen, an denen die Theorie scheitern könnte.

3.6.2 Die Signal-Entdeckungs-Theorie

Einen deutlichen Schritt in diese Richtung unternimmt die Signal-Entdeckungs-Theorie, die zu den unabhängigen und abhängigen Variablen zwar ebenfalls intervenierende Variablen hinzukonstruiert, auf deren operationale Definition jedoch verzichtet.

Treffer- und Fehlalarmquote

Als Beispiel betrachten wir Experimente zum Wiedererkennen (recognition) von zuvor gelernten ("alten") Wörtern in Wörterlisten, die auch andere ("neue") Wörter enthalten, die zuvor nicht gelernt wurden (Klatzky, 1980, 241ff).

Als Trefferquote wird dabei der Anteil der richtig wiedererkannten alten Wörter bezeichnet. Den Anteil der fälschlich für bekannt gehaltenen neuen Wörter bezeichnet man als Fehlalarmquote.

Um das Zustandekommen von Treffern und Fehlalarmen zu erklären, werden der Bekanntheitsgrad der Wörter und die Entscheidungsstrategie der Vpn als intervenierende Variablen eingeführt und die folgenden Axiome postuliert:

Bekanntheitsgrad und Entscheidungsstrategie

1. Jedes Wort hat einen bestimmten Bekanntheitsgrad.
2. Die Bekanntheitsgrade der im Experiment präsentierten Wörter sind normalverteilt.
3. Lernen einer Wörterliste verändert den Bekanntheitsgrad aller in ihr enthaltenen Wörter um denselben Betrag (d').
4. Wörter werden genau dann als bekannt bezeichnet, wenn ihr Bekanntheitsgrad einen bestimmten Kriteriumswert (β) übersteigt.
5. Dieser Kriteriumswert ist sowohl intra- als auch inter-individuell variabel und wird als Entscheidungsstrategie bezeichnet.

Aus den Axiomen 2 und 3 folgt, dass die Bekanntheitsgrade sowohl der alten Wörter (X) als auch der neuen Wörter (Y) normalverteilt sind und sich lediglich die Erwartungswerte der beiden Verteilungen um die Bekanntheitsdifferenz (d') unterscheiden, so dass $E(X) = E(Y) + d'$ (vgl. Abb. 3.6.1). Setzen wir o.B.d.A. voraus, dass die Bekanntheitsgrad-Skala so normiert wurde, dass $E(Y) = 0$ und $\sigma(Y) = 1$, dann folgt für die Fehlalarmquote

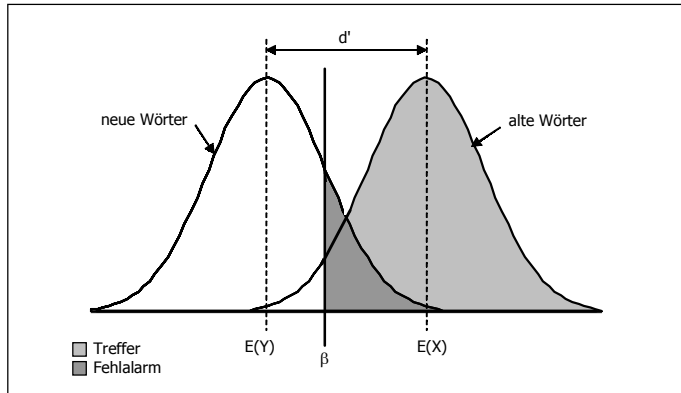
$$\text{prob}\{Y \geq \beta\} = 1 - \Phi(\beta)$$

und für die Trefferquote

$$\text{prob}\{X \geq \beta\} = 1 - \Phi(\beta - d'),$$

worin Φ die Verteilungsfunktion der $N(0,1)$ bezeichnet, wie sie aus jeder Normalverteilungstabelle abgelesen werden kann.

Abb. 3.6.1: Anwendung der Signal-Entdeckungs-Theorie auf das Wiedererkennen von Wörtern



Der tautologische Charakter der Axiome

Aus der Treffer- und Fehlalarmquote einer V_p kann somit jederzeit auf die Bekanntheitsdifferenz zwischen alten und neuen Wörtern (d') und auf die Entscheidungsstrategie (β) geschlossen werden. Umgekehrt kann aus Bekanntheitsdifferenz und Entscheidungsstrategie wieder auf die Treffer- und Fehlalarmquote zurückgerechnet werden.

Da Bekanntheitsdifferenz und Entscheidungsstrategie jedoch nicht unabhängig von Treffer- und Fehlalarmquote gemessen werden können, handelt es sich bei dieser Prognose nicht um eine Erklärung der Treffer und Fehlalarme, sondern schlicht um eine Tautologie. Entsprechend können auch die in den Axiomen

1 bis 5 niedergelegten Modellannahmen der Signal-Entdeckungstheorie (noch) nicht an der Erfahrung scheitern.

Gleichwohl hat die Signal-Entdeckungstheorie jedoch insofern einen gewissen *heuristischen* Wert, als sie den Prozess des Wiedererkennens in zwei getrennte Teilprozesse zerlegt und die Formulierung empirisch gehaltvoller Hypothesen darüber erlaubt, welcher dieser Teilprozesse durch welche Art von unabhängiger Variable beeinflusst wird. So kann man z.B. die Hypothesen formulieren,

Empirische
Prognosen

- dass zusätzliche Lerndurchgänge ausschließlich den Bekanntheitsgrad der Wörter erhöhen (= größeres d'), die Strategie der V_p jedoch unberührt lassen, oder
- dass die Androhung von Strafpunkten für Fehlalarme in einer restriktiveren Strategie resultieren (= größeres β), den Bekanntheitsgrad der Wörter jedoch unberührt lassen.

Aufgrund der Axiome 1 und 3–5 ergeben sich dann die empirischen Prognosen,

- dass zusätzliche Lerndurchgänge zu einer Erhöhung der Trefferquote bei gleichbleibender Fehlalarmquote führen,
- während die Androhung von Strafpunkten für Fehlalarme sowohl die Fehlalarmquote als auch die Trefferquote reduziert.

Die genannten Axiome stellen den strukturellen Kern der Theorie dar, welcher derartige Prognosen ermöglicht, der selbst aber keinen empirischen Gehalt hat und daher auch nicht an der Erfahrung scheitern kann.

Strukturelle und
empirische
Theorieanteile

- Sollte sich eine Prognose als falsch erweisen, so wird dadurch nicht dieser strukturelle Theoriekern falsifiziert, sondern ausschließlich die Hypothese, aus welcher die Prognose abgeleitet wurde.
- Das ist auch dann noch der Fall, wenn sämtliche Hypothesen, die sich auf Grundlage der Axiome sinnvoll formulieren lassen, an der Erfahrung scheitern sollten.

In diesem Falle wird man die Zerlegung des Wiedererkennens in die beiden Teilprozesse des Signals (= Bekanntheitsgrad) und der Strategie (= Kriterium) zwar fallen lassen, aber nicht, weil sie falsifiziert wurde, sondern weil sie sich als nicht zweckmäßig erwiesen hat.

Nimmt man das 2. Axiom (= Normalverteilungsannahme) hinzu, so verschärfen sich die empirischen Prognosen insofern, als sich z.B. aus dem Wechsel zu einer restriktiveren Strategie dann

- nicht nur eine Reduktion sowohl der Fehlalarmquote als auch der Trefferquote ableiten lässt,
- sondern auch die genauen Zahlenverhältnisse prognostiziert werden können, in denen sich die beiden Quoten verändern.¹

Diese Zahlenverhältnisse lassen sich ganz einfach aus der Tabelle für die Standardnormalverteilung ablesen. So entspricht z.B. einer Fehlalarmquote von 0.94% ein Kriterium von $\beta = 2.35$. Bei einer Trefferquote von 63.68% ergibt sich daraus eine Bekanntheitsdifferenz von $d' = 2.7$, und bei einer Verschärfung des Kriteriums müssten sich Fehlalarm- und Trefferquote so verhalten wie in Tab. 3.6.1 dargestellt:

Tab. 3.6.1: Veränderung der Fehlalarm- und Trefferquote bei Verschärfung des Kriteriums

β	$d' = 2.7$	
	Fehlalarmquote	Trefferquote
2.35	0.94%	63.68%
2.40	0.82%	61.79%
2.45	0.71%	59.87%
2.50	0.62%	57.93%

Im Kontext von Hypothesen, welche Strategieänderungen bei gleichbleibender Bekanntheitsdifferenz zum Gegenstand haben, erweist sich somit auch die in Axiom 2 niedergelegte Normalverteilungsannahme als empirisch gehaltvoll.

Gerade diese Annahme ist aber völlig beliebig² und rein technischer Natur. Sie dient lediglich dazu, die intervenierenden Variablen d' und β quantitativ bestimmen zu können. Für die Erklärung des Wiedererkennens durch die intervenierenden Variablen ist sie verzichtbar.

1. Dasselbe gilt auch für den Wechsel zu einer weniger restriktiven Strategie, die eine gleichzeitige Erhöhung der beiden Quoten nach sich zieht.
 2. Die Normalverteilungsannahme entbehrt jeder Begründung, und an Stelle der Normalverteilung könnte genau so gut jede andere Verteilungsform postuliert werden.

3.6.3 Sieglers Theorie der kognitiven Entwicklung im Kindesalter

Ganz im Sinne des Neobehaviorismus erklären die Informationsverarbeitungsmodelle das Verhalten des Systems, indem sie Annahmen über die in der black box ablaufenden Prozesse treffen. Zugleich wird das behavioristische Paradigma damit aber insofern verlassen, als die empirischen Prognosen nicht mehr (oder zumindest nicht primär) auf dem Ursache-Wirkungs-Prinzip beruhen, sondern nach Art einer intentionalen Erklärung konstruiert werden:

Intentionale
Erklärung

Das System verfolgt das Ziel, auf eingehende Informationen (Input) angemessen zu reagieren (Output).

Als Mittel zu diesem Zweck dienen dem System bestimmte Informationsverarbeitungsstrategien.

Also produziert das System jenen Output, der aus diesen Informationsverarbeitungsstrategien resultiert.

Empirische Hypothesen kommen erst (sekundär) ins Spiel, wenn Annahmen darüber getroffen werden, welche Strategien unter welchen Bedingungen zum Einsatz kommen.

Vielleicht noch deutlicher als an der oben diskutierten Signal-Entdeckungs-Theorie wird dies am Beispiel von Sieglers Theorie der kognitiven Entwicklung im Kindesalter (Siegler, 1976; Klahr & Siegler, 1978; Klahr 1978), die sich im Vergleich zur Signal-Entdeckungs-Theorie durch eine umfassende handlungstheoretische Begründungsfähigkeit – sowohl des Modells als auch der daraus resultierenden empirischen Prognosen – sowie durch den Verzicht auf empirisch gehaltvolle, gleichwohl aber bloß technische Zusatzannahmen auszeichnet.

Handlungstheoretische
Begründungsfähigkeit
und Verzicht auf
technische
Zusatzannahmen

So erlaubt Sieglers Theorie nicht nur (wie die Signal-Entdeckungs-Theorie auch) eine empirische Reduktion der Informationsverarbeitungsstrategien auf das Antwortverhalten der Vpn (vgl. Kapitel 2.4.7), sondern auch deren genetische Rekonstruktion und pragmatische Evaluation.

Die von Siegler unterschiedenen Entwicklungsstufen enthalten eine bestimmte Entwicklungslogik, wobei jede Stufe auf der vorangegangenen aufbaut und die für eine Stufe typische Informationsverarbeitungsstrategie aus der jeweils vorangegangenen entsteht, indem weiteres Wissen über die physikalischen Gesetz-

Genetische
Rekonstruktion

mäßigkeiten integriert wird, welche das Verhalten der Balkenwaage bestimmen.

- Zuerst wissen die Kinder lediglich, dass das Verhalten der Balkenwaage etwas mit Gewichten zu tun hat. Das größere Gewicht entscheidet (Stufe 1).
- Dann lernen die Kinder, dass sich die Waage auch bei gleichen Gewichten nicht immer in Balance befindet, sondern dass in diesem Fall die größere Distanz entscheidend wird (Stufe 2).
- Haben sie erst einmal erkannt, dass die Distanzen ebenfalls eine Rolle spielen, so können sie in einem nächsten Schritt lernen, dass dies nicht nur bei gleichen Gewichten der Fall ist. Da sie aber noch nicht wissen, wie Gewicht und Distanz miteinander zu verrechnen sind, ist dies zunächst verwirrend, und die Kinder können bei den Konflikt-Aufgaben nur raten (Stufe 3).
- Am Ende haben sie den richtigen Algorithmus gefunden und bewältigen auch die Konflikt-Aufgaben, indem sie das Drehmoment (= Produkt aus Gewicht x Distanz) berechnen (Stufe 4).

Typische
Entwicklungsverläufe

Ersichtlich ist diese genetische Rekonstruktion des *typischen* Entwicklungsverlaufes aber nicht nur eine Frage der Sachlogik, sondern auch eine empirische Frage: d.h. sie ergibt sich nicht nur daraus, welche Erfahrungen Kinder machen können, wenn sie ein bestimmtes Wissen bereits haben, sondern auch daraus, welche Erfahrungen sie *normalerweise* tatsächlich machen.

Tab. 3.6.2: Genetische Rekonstruktion des Wechsels von Stufe 1 auf Stufe 2 nach dem narrativen Erklärungsmodell

t ₁	Das Kind befindet sich auf Stufe 1 und beachtet lediglich die Gewichte.
t ₂	Das Kind macht die Erfahrung, dass sich die Balkenwaage bei Aufgaben vom Typ Distanz trotz gleicher Gewichte nicht in Balance befindet. Das Kind sucht nach der Ursache, warum sich die Balkenwaage bei dieser Art von Aufgaben anders verhält als sie sollte. Das Kind bemerkt, dass sich bei diesen Aufgaben auch die Distanzen unterscheiden. Das Kind vermutet, dass die unterschiedlichen Distanzen das Verhalten der Balkenwaage beeinflussen könnten. Das Kind macht die Erfahrung, dass bei dieser Art von Aufgaben immer die Seite mit der größeren Distanz nach unten geht. Das Kind integriert diese Erfahrung in seine Informationsverarbeitungsstrategie.
t ₃	Das Kind befindet sich auf Stufe 2 und beachtet bei gleichen Gewichten auch die Distanzen.

Um z.B. zu erklären, dass ein Kind, das sich kürzlich (t_1) noch auf Stufe 1 befunden hat, sich jetzt (t_3) auf Stufe 2 befindet, könnten wir etwa rekonstruieren, dass das Kind in der Zwischenzeit (t_2) die Erfahrung gemacht hat, dass sich die Balkenwaage bei Aufgaben vom Typ ‚Distanz‘ trotz gleicher Gewichte nicht in Balance befindet, dass es nach einer Ursache dafür gesucht und dabei bemerkt hat, dass sich bei diesen Aufgaben auch die Distanzen unterscheiden; dass es vermutet hat, dies könnte die Ursache sein; dass es dann die Erfahrung gemacht hat, dass diese Vermutung richtig war; und dass es dieses neu erworbene Wissen schließlich in seine Informationsverarbeitungsstrategie integriert hat (vgl. Tab. 3.6.2).

Wenn wir von einem typischen Entwicklungsverlauf sprechen, dann sagen wir damit nichts anderes, als dass Kinder (unter den gegebenen kulturellen Rahmenbedingungen) normalerweise früher oder später diese (oder gleichwertige) Erfahrungen machen.

In ähnlicher Weise kann man auch den Wechsel von Stufe 2 auf Stufe 3 und von Stufe 3 auf Stufe 4 genetisch rekonstruieren. Dabei zeigt sich, dass von Stufe 3 auf Stufe 4 ein unvermittelter Sprung stattfindet:

- Auf Stufe 3 werden die Kinder von den Konflikt-Aufgaben noch verwirrt, weil sie keine Ahnung haben, wie man Gewicht und Distanz miteinander verrechnen könnte.
- Auf Stufe 4 haben sie bereits den korrekten Lösungsalgorithmus gefunden und bewältigen die Konflikt-Aufgaben durch Berechnung des Drehmoments.

Die Erweiterung von Sieglers Theorie durch May (1979)

t_1	Das Kind befindet sich auf Stufe 3 und kann bei den Konflikt-Aufgaben nur raten.
t_2	Das Kind gibt sich damit nicht zufrieden und sucht nach einem Algorithmus, wie Gewicht und Distanz bei den Konflikt-Aufgaben miteinander zu verrechnen sind. Das Kind vermutet, dass das Produkt aus Gewicht \times Distanz für das Verhalten der Balkenwaageentscheidend sein könnte. Das Kind berechnet bei den Konflikt-Aufgaben das Produkt aus Gewicht \times Distanz und macht die Erfahrung, dass es das Verhalten der Balkenwaage damit richtig vorhersagen kann. Das Kind integriert diese Erfahrung in seine Informationsverarbeitungsstrategie.
t_3	Das Kind befindet sich auf Stufe 4 und löst die Konfliktaufgaben durch Berechnung des Drehmomentes.

Tab. 3.6.3: Genetische Rekonstruktion des Wechsels von Stufe 3 auf Stufe 4 nach dem narrativen Erklärungsmodell

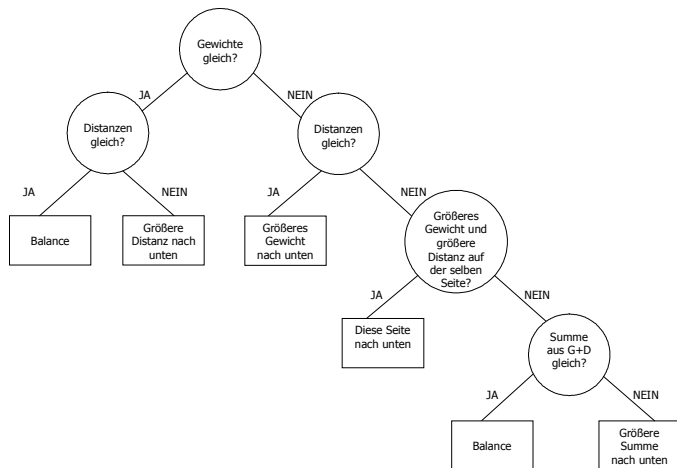
Zur genetischen Rekonstruktion des Wechsels von Stufe 3 auf Stufe 4 (vgl. Tab. 3.6.3) muss man entsprechend voraussetzen, dass die Kinder entweder auf Anhieb den richtigen Algorithmus finden, oder so lange im Zustand der Verwirrung und des bloßen Ratens bleiben, bis sie ihn gefunden haben.

Plausibler wäre es, anzunehmen, dass die Kinder es zunächst mit der einfachsten Rechenart versuchen und an Stelle des Produktes die Summe aus Gewicht + Distanz berechnen (May, 1979). Sollte dies der Fall sein, so gibt es zwei Möglichkeiten:

- Die Kinder versuchen es halt ein paar Mal; bemerken, dass dies nicht die Lösung ist und bleiben im Zustand der Verwirrung und des Ratens, oder
- die Kinder bemerken, dass sie auf diese Weise zumindest manche der Konflikt-Aufgaben lösen können und bedienen sich dieses Algorithmus, obwohl er fehlerhaft ist.

Wenn es zumindest einige Kinder gibt, welche die letztgenannte Möglichkeit ergreifen, dann müsste es zwischen Stufe 3 und Stufe 4 noch eine Zwischenstufe 3A geben, die sich der in Abb. 3.6.2 dargestellten Lösungsstrategie bedient.

Abb. 3.6.2: Die Entwicklungsstufe 3A nach May (1979)



Empirische Reduktion Um die Entwicklungsstufen 3 und 3A auch empirisch unterscheiden zu können, reichen die von Siegler entwickelten Aufgabentypen allerdings nicht aus. Vor allem die Konflikt-Aufgaben müssen weiter differenziert werden (vgl. Tab. 3.6.4). Hatte Siegler je

nach richtiger Antwort (Stufe 4) drei Typen von Konflikt-Aufgaben unterschieden, so sind es bei May zehn, wobei z.B. Konflikt-Gewicht/Distanz einen Typ von Konflikt-Aufgaben bezeichnet, bei dem sich die Kinder nach der Produktregel (Stufe 4) für die Seite mit dem größeren Gewicht, nach der Summenregel (Stufe 3A) dagegen für die Seite mit der größeren Distanz entscheiden. Als Harmonie-Aufgaben bezeichnet May jene Konflikt-Aufgaben, bei denen beide Regeln zum selben Ergebnis führen.

Aufgabentyp	Beispiel				Stufe				
	Links		Rechts		1	2	3	3A	4
	G	D	G	D					
Balance	2	3	2	3	B	B	B	B	B
Gewicht	1	1	5	1	G	G	G	G	G
Distanz	3	2	3	3	B	D	D	D	D
Harmonie	4	2	3	1	G=D	G=D	G=D	G=D	G=D
Konflikt-Gewicht/Gewicht	1	3	4	1	G	G	?	G	G
Konflikt-Gewicht/Distanz	1	6	3	3	G	G	?	D	G
Konflikt-Gewicht/Balance	3	6	5	4	G	G	?	B	G
Konflikt-Distanz/Gewicht	3	3	6	1	G	G	?	G	D
Konflikt-Distanz/Distanz	1	4	3	1	G	G	?	D	D
Konflikt-Distanz/Balance	4	5	6	3	G	G	?	B	D
Konflikt-Balance/Gewicht	2	2	4	1	G	G	?	G	B
Konflikt-Balance/Distanz	1	4	2	2	G	G	?	D	B
Konflikt-Balance/Balance	2	3	3	2	G	G	?	B	B
Erwartete Anzahl richtiger Antworten ($p=0.33$)					6	7	7	7	13

Tab. 3.6.4: Prognostizierte Antwortmuster für 13 Aufgabentypen und 5 Entwicklungsstufen. Bezeichnung der Konflikt-Aufgaben: Stufe 4/Stufe 3A (Produktregel/Summenregel)

Wie May an einer Stichprobe von $n = 445$ Schülern der 5. Klasse des Gymnasiums¹ zeigt, gibt es die von ihr postulierte Stufe 3A tatsächlich. May vermutet sogar, "daß das von Siegler (1976) beschriebene ‚muddle through‘ tatsächlich eine Anwendung der Summenregel darstellt, die sich bei Siegler der Beobachtung entzog, da (...) seine Aufgaben nicht dazu geeignet waren, zwischen der Anwendung einer Additionsregel und einem Rateverhalten zu differenzieren" (May, 1979, 71).

Mangels Verfügbarkeit leistungsfähiger Computerprogramme zur Berechnung der Latent-Class-Analyse musste sich May bei der Auswertung ihrer Daten allerdings noch mit Näherungsverfahren

1. Durchschnittsalter 11 Jahre; 41% weiblich, 59% männlich.

begnügen. Die Identifizierbarkeit der Entwicklungsstufen mittels LCA wurde erst von Kolb (1993) mittels Simulationsstudien nachgewiesen,¹ und Balogh (1994) zeigte schließlich in einer Testuntersuchung an 163 Schülern der 5. und 6. Klasse des Gymnasiums,² dass die Unterscheidung der Entwicklungsstufen mittels LCA auch in empirischem Datenmaterial (Papier-Bleistift Test) gut gelingt. Die von Siegler beschriebene Entwicklungsstufe 3 wurde auch von ihr nicht gefunden, was für die von May geäußerte Vermutung sprechen könnte; zumindest aber dafür, dass es sich hierbei allenfalls um ein kurzes Übergangsstadium handelt.

Pragmatische Evaluation

Ein grundlegendes Charakteristikum der Theorie von Siegler besteht darin, dass sie nicht nur den kognitiven Entwicklungsstand der Kinder als mehr oder weniger weit fortgeschritten beschreibt, sondern präzise Aussagen darüber macht, wie die Kinder an die Problemlösung herangehen, welche Informationen sie aufnehmen und wie sie diese verarbeiten. Auf dieser Grundlage werden dann konkrete Handlungsanweisungen möglich, wie die Kinder in ihrer Entwicklung gezielt gefördert werden können: indem man

1. an ihren gegenwärtigen Wissensstand anknüpft und sie
2. mit jenem Typ von Aufgaben konfrontiert, deren richtige Beantwortung auf ihrer gegenwärtigen Entwicklungsstufe noch nicht gelingt, sondern erst auf der nächst höheren Stufe;
3. indem man ihnen ein Feedback über das tatsächliche Verhalten der Waage gibt und die Kinder dafür
4. am besten mit echten Balkenwaagen spielen lässt, statt sie lediglich mit Papier-Bleistift-Tests und verbalen Feedbacks zu konfrontieren.

Schon die empirische Reduktion der Entwicklungsstufen setzt ja voraus, dass die Kinder das Verhalten der Balkenwaage auch wirklich vorhersagen wollen; und auch für die Lernprozesse, welche durch die Fördermaßnahmen angestoßen werden sollen, dürften motivationale Faktoren nicht gerade unbedeutend sein.

-
1. Allerdings mit der Einschränkung, dass die Unterscheidung zwischen Stufe 1 und Stufe 2 mittels der oben dargestellten 13 Aufgaben nicht zuverlässig möglich ist, weil der Unterschied zwischen den beiden Stufen nur bei einem Aufgabentyp (Distanz-Aufgaben) zum Tragen kommt. Um eine zuverlässige Unterscheidung zu ermöglichen, sollte man bei der Testkonstruktion daher mehrere Aufgaben dieses Typs vorsehen.
 2. Durchschnittsalter 11 Jahre 4 Monate; 57% weiblich, 43% männlich.

3.6.4 Die statistische Regelmäßigkeit des Entwicklungsverlaufes

Über die Identifikation der Informationsverarbeitungsstrategien und die Diagnose der Entwicklungsstufen hinaus können aus der Theorie von Siegler eine Reihe von entwicklungspsychologischen Regelmäßigkeiten abgeleitet werden.

Dabei handelt es sich um *empirische* Regelmäßigkeiten, die aber nicht Naturgesetze im Sinne des deduktiv-nomologischen Erklärungsmodells darstellen, für deren Geltung es keinen logischen oder sachlogischen Grund gibt (vgl. Kapitel 2.3); sondern um statistische Regelmäßigkeiten, die sich aufgrund *struktureller* Gesetzmäßigkeiten aus den empirischen Randbedingungen ergeben.

Als strukturelle Gesetzmäßigkeiten kommen dabei einerseits das intentionale Erklärungsmodell und andererseits die Entwicklungslogik der Informationsverarbeitungsstrategien zum Tragen. Empirische Randbedingungen sind die Motivation der Probanden; das kulturelle Umfeld, das es ihnen erlaubt (früher oder später), bestimmte (typische) Erfahrungen zu machen, sowie die Integration dieser Erfahrungen in die eigenen Informationsverarbeitungsstrategien.

1. Aufgrund der beschriebenen Entwicklungslogik ist davon auszugehen, dass die Entwicklungsstufen alterskorreliert sind – auch wenn manche Kinder die jeweils nächste Stufe etwas früher erreichen mögen als andere und selbst dann, wenn die höchste Entwicklungsstufe nur von einem Bruchteil der Kinder überhaupt je erreicht werden sollte.
2. Mit zunehmendem Lebensalter bleibt die Leistungsmenge über lange Strecken konstant und macht dann einen plötzlichen Sprung (vgl. Abb. 3.6.3).¹

1. Der genaue Entwicklungsverlauf wird natürlich von der jeweiligen Verweildauer auf den verschiedenen Entwicklungsstufen abhängen. Die Abbildungen 3.6.3-3.6.8 beruhen auf den vereinfachenden Annahmen, dass die Verweildauer auf den Stufen 1, 2 und 3+3A gleich lang ist (jeweils 3 Einheiten), dass zwischen den einzelnen Stufen eine Übergangsphase (2 Einheiten) stattfindet, in der zunächst 25%, dann 75% der Kinder bereits die nächsthöhere Stufe erreicht haben. Im Falle der Stufe 3+3A findet zunächst der Übergang von Stufe 2 auf Stufe 3 und sofort danach (ohne Verweilen auf Stufe 3) der Übergang von Stufe 3 auf Stufe 3A. Die Verweildauer auf Stufe 3A verkürzt sich dadurch auf 2 Einheiten.

1

Abb. 3.6.3: Zunahme der Leistungsmenge mit wachsendem Lebensalter¹

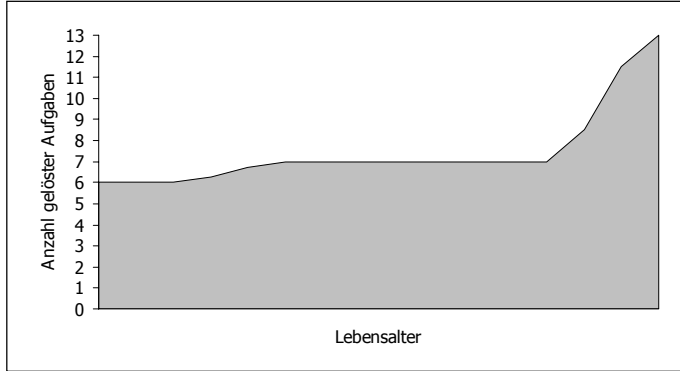


Abb. 3.6.4: Leistungsentwicklung bei den Distanzaufgaben

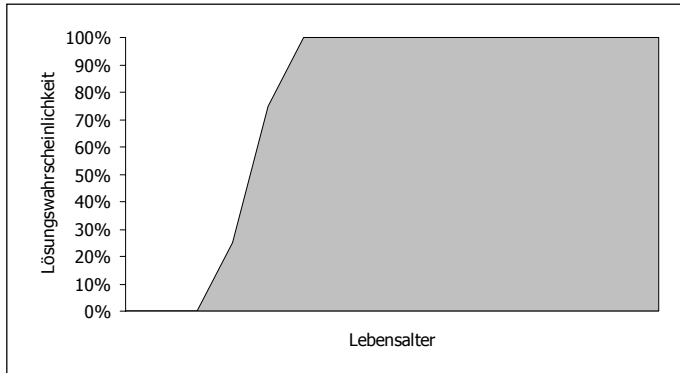
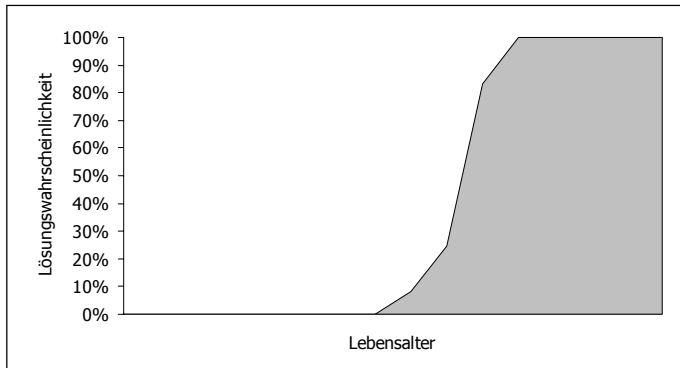


Abb. 3.6.5: Leistungsentwicklung bei Aufgaben vom Typ ‚Konflikt-Distanz/Distanz‘ und ‚Konflikt-Balance/Balance‘



1. Aus Gründen der Einfachheit gehen wir in Abb. 3.6.3 von je einer Aufgabe pro Aufgabentyp aus.

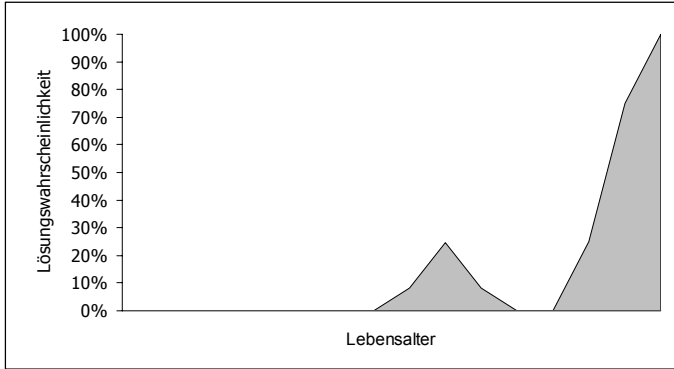


Abb. 3.6.6: Leistungsentwicklung bei Aufgaben vom Typ ‚Konflikt-Distanz/Gewicht‘, ‚Konflikt-Distanz/Balance‘, ‚Konflikt-Balance/Gewicht‘ und ‚Konflikt-Balance/Distanz‘

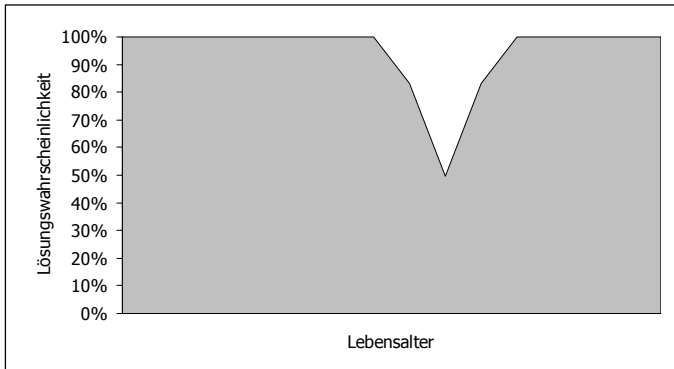


Abb. 3.6.7: Leistungsentwicklung bei Aufgaben vom Typ ‚Konflikt-Gewicht/Gewicht‘

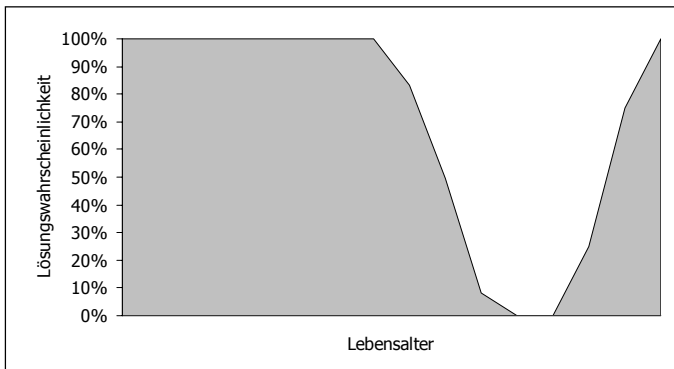


Abb. 3.6.8: Leistungsentwicklung bei Aufgaben vom Typ ‚Konflikt-Gewicht/Distanz‘ und ‚Konflikt-Gewicht/Balance‘

3. Bei Balance- Aufgaben, Gewicht-Aufgaben und Harmonie-Aufgaben findet über die Zeitachse hinweg keine Leistungsveränderung statt: Die Kinder beantworten diese Aufgaben von Anfang an richtig (d.h. sobald sie überhaupt mit den Balkenwaagen-Aufgaben zurechtkommen).
4. Bei Distanz-Aufgaben gibt es von Stufe 1 auf Stufe 2 einen sprunghaften Leistungsanstieg (vgl. Abb. 3.6.4).
5. Bei Aufgaben vom Typ ‚Konflikt-Distanz/Distanz‘ und ‚Konflikt-Balance/Balance‘ steigt die Leistung dagegen erst von Stufe 2 bis Stufe 3A an (vgl. Abb. 3.6.5).
6. Bei Aufgaben vom Typ ‚Konflikt-Distanz/Gewicht‘, ‚Konflikt-Distanz/Balance‘, ‚Konflikt-Balance/Gewicht‘ und ‚Konflikt-Balance/Distanz‘ findet auf Stufe 3 zwar ebenfalls eine geringfügige Leistungsverbesserung statt. Diese geht auf Stufe 3A jedoch wieder verloren. Erst von Stufe 3A auf Stufe 4 ergibt sich dann ein plötzlicher Leistungssprung (vgl. Abb. 3.6.6).
7. Aufgaben vom Typ ‚Konflikt-Gewicht/Gewicht‘ werden von Anfang an richtig beantwortet. Auf Stufe 3 gibt es jedoch einen vorübergehenden Leistungseinbruch (vgl. Abb. 3.6.7).
8. Aufgaben vom Typ ‚Konflikt-Gewicht/Distanz‘ und ‚Konflikt-Gewicht/Balance‘ werden zunächst richtig beantwortet. Von Stufe 2 bis Stufe 3A fällt die Leistung jedoch kontinuierlich ab. Von Stufe 3A auf Stufe 4 steigt sie wieder sprunghaft an (vgl. Abb. 3.6.8).

Die
Erklärungsfähigkeit
statistischer
Regelmäßigkeiten

Die statistischen Regelmäßigkeiten der Leistungsentwicklung, wie wir sie oben dargestellt haben, können zwar herangezogen werden, um auch im konkreten Fall *plausibel* zu machen, warum die Leistungsentwicklung bei bestimmten Aufgabentypen nicht linear verläuft. Dass sich z.B. die Leistung eines Kindes bei Aufgaben vom Typ ‚Konflikt-Gewicht/Distanz‘ oder ‚Konflikt-Gewicht/Balance‘ im Laufe seiner Entwicklung (vorübergehend) verschlechtert, ist bei Kenntnis der in Abb. 3.6.8 dargestellten statistischen Regel erwartbar bzw. kann durch die Angabe dieser Regel induktiv gestützt werden. Wie die Theorie von Siegler zeigt, muss man sich mit einer solchen statistisch-induktiven Erklärung aber nicht schon zufrieden geben, sondern statistische Regelmäßigkeiten können ihrerseits als erklärungsbedürftiger Zusammenhang aufgefasst und – im gegebenen Fall durch den sukzessiven Wissenszuwachs der Kinder – erklärt werden.

Kontrollfragen

- Worin besteht der Unterschied zwischen Behaviorismus und Neobehaviorismus?
- Wie lauten die Grundannahmen der kognitiven Lerntheorie von Tolman?
- Wie definiert Tolman die in seiner Theorie enthaltenen intervenierenden Variablen?
- Worauf gründet sich der Verdacht, dass es sich bei der Theorie von Tolman um eine Verdoppelung der Realität handeln könnte?
- Wie könnte die (vermutete) Realitätsverdoppelung behoben werden?
- Wie sind in der Signal-Entdeckungs-Theorie die folgenden Termini definiert?
 - a) Trefferquote,
 - b) Fehlalarmquote,
 - c) Bekanntheitsgrad
 - d) Entscheidungsstrategie?
- Wie lauten die Axiome der Signal-Entdeckungs-Theorie?
- Inwiefern handelt es sich bei der Prognose der Treffer- und Fehlalarmquote um eine Tautologie?
- Worin besteht der heuristische Wert der Signal-Entdeckungs-Theorie?
- Geben Sie Beispiele für empirisch gehaltvolle Hypothesen bezüglich der Veränderung
 - a) des Bekanntheitsgrades der Wörter,
 - b) der Entscheidungsstrategie einer Vpund zeigen Sie, wie sich Treffer- und Fehlalarmquote bei Geltung dieser Hypothesen verhalten müssen!
- Erläutern Sie, welche der Axiome den strukturellen Kern der Signal-Entdeckungs-Theorie bilden, welches der Axiome lediglich eine technische Zusatzannahme darstellt, und welche der Axiome empirisch gehaltvoll sind!
- Diskutieren Sie, inwiefern die Signal-Entdeckungs-Theorie an der Erfahrung scheitern kann!
- Erläutern Sie, inwiefern Informationsverarbeitungsmodelle
 - a) zwar einerseits dem neobehavioristischen Ansatz entsprechen,
 - b) zugleich aber das behavioristische Paradigma verlassen!
- Konkretisieren Sie das intentionale Erklärungsmodell, welches den Informationsverarbeitungsmodellen zugrunde liegt!
- Erläutern Sie die Entwicklungslogik, die Sieglers Entwicklungsstufen zu eigen ist!
- Erläutern Sie, inwiefern sich die genetische Rekonstruktion typischer Entwicklungsverläufe auf strukturelle Gesetzmäßigkeiten stützen kann und inwiefern dabei empirische Faktoren zum Tragen kommen!
- Erläutern Sie die Behauptung, dass ein Entwicklungsverlauf typisch ist!
- Stellen Sie die genetische Rekonstruktion des Wechsels von Stufe 2 auf Stufe 3 im narrativen Erklärungsmodell dar!
- Erläutern Sie,
 - a) welche Erweiterung der Siegler'schen Theorie durch May (1979)

vorgenommen wurde,

b) welche Konsequenzen sich daraus für die empirische Reduktion der Entwicklungsstufen ergeben, und

c) durch welche empirischen Befunde Mays Erweiterung der Theorie gestützt wird!

- Welche Fördermaßnahmen würden Sie vorsehen, um ein Kind, das sich
 - a) auf Stufe 1,
 - b) auf Stufe 2,
 - c) auf Stufe 3 bzw.
 - d) auf Stufe 3Abefindet, an die nächst höhere Entwicklungsstufe heranzuführen?
- Wie kann mittels der Entwicklungsstufen nach Siegler bzw. May erklärt werden,
 - a) dass im Laufe der Entwicklung ein plötzlicher Leistungssprung stattfindet, nachdem die Leistungsmenge zuvor über lange Zeit konstant blieb, sowie
 - b) dass nicht bei allen Typen von Aufgaben ein monotoner Leistungsanstieg zu verzeichnen ist?
- Erläutern Sie,
 - a) inwiefern sich diese Verlaufsgesetzmäßigkeiten von Naturgesetzen im Sinne des deduktiv-nomologischen Erklärungsmodells unterscheiden sowie
 - b) auf welchen empirischen und strukturellen Prämissen sie beruhen!
- Diskutieren Sie am Beispiel der Theorie von Siegler bzw. ihrer Erweiterung durch May, inwiefern statistische Regelmäßigkeiten bereits eine Erklärung des jeweiligen Sachverhalts darstellen oder als Beschreibung eines erklärungsbedürftigen Zusammenhanges aufgefasst werden können!

3.7

Strukturelle und empirische Theorieanteile

3.7.1 Der Zirkularitätsverdacht gegenüber dem Gesetz des Effektes

Der Verdacht, dass so manche psychologische Theorie eine Tautologie sein könnte, ist schon relativ alt. So hat bereits Postman (1947) den Vorwurf erhoben, dass Thorndykes (1913) Gesetz des Effektes schlichtweg zirkulär sei.

Ausgangspunkt des Gesetzes des Effektes ist bekanntlich, dass Organismen auf neue Problemsituationen zunächst so lange probeweise mit instinktiven oder bereits früher erlernten Verhaltensweisen reagieren, bis eine Reaktion zufällig zum Erfolg führt.

Das Gesetz des Effektes

Thorndykes Gesetz besagt dann, dass Verknüpfungen von Situationen und Reaktionen, die von einem befriedigenden Gesamtzustand (satisfying state of affairs) begleitet sind, verstärkt werden, während die Stärke von Verknüpfungen, die von einem unbefriedigenden bzw. unangenehmen Gesamtzustand (annoying state of affairs) gefolgt sind, abnimmt (vgl. Kapitel 1.1.2).

Damit wird eine Erklärung des Lernens nach dem deduktiv-nomologischen Erklärungsmodell angestrebt. Das Gesetz des Effektes spielt darin die Rolle der Gesetzesaussage und hat die logische Struktur:

$$\forall x. A(x) \rightarrow B(x)$$

"Für alle Verknüpfungen gilt, dass sie verstärkt werden, falls sie von einem befriedigenden Gesamtzustand gefolgt werden, bzw. dass ihre Stärke abnimmt, falls sie von einem unbefriedigenden Gesamtzustand gefolgt sind."

Damit eine Deduktion nach dem HO-Schema aber tatsächlich als eine Erklärung gelten kann, müssen einige weitere Voraussetzungen erfüllt sein; u.a., dass im Antezedens (UV) und Succedens (AV) der Gesetzesaussage ausschließlich empirische Aussagen vorkommen, sowie, dass die Aussagen, welche den zu erklären-

Adäquatheitsbedingungen der deduktiv-nomologischen Erklärung

den Sachverhalt darstellen (= positive bzw. negative Verstärkung der Verknüpfung), nicht bereits (sach-)logisch aus den Aussagen herleitbar sein dürfen, welche die Antezedensbedingungen (= befriedigender bzw. unbefriedigender Gesamtzustand) darstellen.

Der
Zirkularitätsvorwurf

Der von Postman erhobene Vorwurf der Zirkularität des Gesetzes des Effektes besagt nun, dass genau diese Voraussetzung nicht erfüllt sei. Die Termini ‚befriedigender‘ bzw. ‚unbefriedigender Gesamtzustand‘ könnten nicht unabhängig von der positiven bzw. negativen Verstärkung der Verknüpfung von Situation und Reaktion definiert werden. Das Gesetz besage daher nicht mehr als: "Das Tier tut das, was es tut, weil es das tut, und es unterlässt das, was es nicht tut, weil es das nicht tut" (Postman, 1947, 96; deutsch nach Foppa, 1965, 335).

Thorndykes Formulierung des Gesetzes wird dieser Vorwurf allerdings insofern nicht ganz gerecht, als Thorndyke in der Tat eine Definition versucht hat, die *nicht* von der verstärkenden Wirkung des befriedigenden Gesamtzustandes Gebrauch macht: Als satisfying state of affairs sollen laut Thorndyke solche Bedingungen gelten, bei denen "das Tier nichts unternimmt, um sie zu vermeiden, aber oft bestrebt ist, sie zu erreichen oder zu erhalten".

Andererseits verwendet Thorndyke in dieser Formulierung aber eine Reihe von intentionalen Termini, die über das Beobachtbare hinausgehen, die daher für den Behaviorismus als unzulässig galten und deren Anwendbarkeit auf das Verhalten von Tieren – auch über die behavioristische Selbstbeschränkung auf das beobachtbare Verhalten hinaus – zumindest zweifelhaft ist.

Zur Überprüfung des Zirkularitätsvorwurfes werden wir folglich in zwei Schritten vorgehen müssen:

- Zunächst ist zu untersuchen, ob eine zirkelfreie Formulierung des Gesetzes innerhalb des Behaviorismus möglich ist. Hier werden wir sehen, dass dies tatsächlich nicht geht.
- Darauf aufbauend, werden wir untersuchen, ob eine zirkelfreie Formulierung des Gesetzes jenseits des Behaviorismus möglich ist. Hier werden wir auf das Prinzip der strukturellen Reduktion zurückkommen und sehen, dass eine zirkelfreie Formulierung des Gesetzes tatsächlich möglich ist, sofern wir die Semantik von Stimuli und Reaktionen beachten.

Für eine konsequent behavioristische Formulierung des Gesetzes stehen zwei Möglichkeiten offen: C.L. Hull

Erstens könnte man in Übereinstimmung mit den meisten experimentellen Anordnungen statt von einem satisfying state of affairs schlicht von einer Triebreduktion sprechen: So werden S-R-Verbindungen bei nahrungs-deprivierten Tieren durch die Gabe von Futterpillen verstärkt.

Dieser Weg, der von C.L. Hull in seinen *Principles of Behavior* (1943) eingeschlagen wurde, scheidet jedoch daran, dass man hierfür, um im Rahmen des Behaviorismus zu bleiben, eine rein beobachtungssprachliche Definition des Terminus ‚Triebreduktion‘ brauchte.

Anders ausgedrückt: Die Triebreduktion müsste durch physiologische Messnormen operationalisierbar sein. Dabei ist jedoch der Wunsch der Vater des Gedanken. Bei allen Fortschritten der Physiologie gilt immer noch der von Meehl (1950, 56f) geäußerte Zweifel: "... that anyone would be able to give even the vaguest specification of the defining property of all neural events which are reinforcing. Even for the so-called-primary physiological needs such as hunger, the evidence as to their exact physiological basis is most incomplete."

Und selbst wenn man das könnte, steht der Hull'schen Formulierung des Gesetzes entgegen, dass die experimentelle behavioristische Forschung bis dato eben *nicht* von solchen physiologischen Messnormen Gebrauch macht, um die Triebreduktion zu messen.

Der andere Weg wurde von Skinner (1953) gewählt. Statt von satisfying und annoying states of affairs spricht Skinner von Belohnung (reward) und Bestrafung (punishment). Die Stärke einer S-R-Verbindung ersetzt Skinner durch die Auftrittswahrscheinlichkeit der Reaktion. B.F. Skinner

Das Gesetz des Effektes wird somit zum Gesetz des operanten Konditionierens:

- Die Auftrittswahrscheinlichkeit einer Reaktion wird erhöht, wenn ihr eine Belohnung folgt, bzw. sie wird verringert, wenn ihr eine Bestrafung folgt.
- Den Terminus ‚Belohnung‘ definiert Skinner als Hinzufügen eines positiven oder Wegnahme eines negativen Verstärkers. Entsprechend die ‚Bestrafung‘ als Wegnahme eines positiven

oder Hinzufügen eines negativen Verstärkers.

- Einen positiven Verstärker definiert Skinner als einen Stimulus, dessen Hinzufügen zu einer Situation die Auftrittswahrscheinlichkeit einer Reaktion verstärkt. Einen negativen Verstärker definiert er als einen Stimulus, dessen Wegnahme aus einer Situation die Auftrittswahrscheinlichkeit einer Reaktion erhöht.

Damit reduziert sich das Gesetz auf die Aussagen,

- dass sich die Auftrittswahrscheinlichkeit einer Reaktion erhöht, wenn ihr eine Stimulusveränderung folgt, welche die Auftrittswahrscheinlichkeit einer Reaktion erhöht, bzw.
- dass sich die Auftrittswahrscheinlichkeit einer Reaktion verringert, wenn ihr eine Stimulusveränderung folgt, welche die Auftrittswahrscheinlichkeit einer Reaktion verringert.¹

In dieser Formulierung ist das Gesetz des Effektes jedenfalls dann zirkulär, wenn beide Male dieselbe Reaktion gemeint ist.

Werden im Succedens und im Antecedens der Gesetzesaussage dagegen verschiedene Reaktionen eingesetzt, so behauptet man auf einmal, dass sich die Auftrittswahrscheinlichkeit einer Reaktion erhöht, wenn ihr eine Stimulusveränderung folgt, welche die Auftrittswahrscheinlichkeit (irgend-)einer anderen Reaktion erhöht. Anders ausgedrückt: dass jede Stimulusveränderung, die für eine Reaktion ein Verstärker ist, auch alle anderen (lernbaren) Reaktionen verstärkt, d.h. dass alle Verstärker transsituational sind.

Dieses von Meehl (1950) so genannte schwache Gesetz des Effektes hat aber mit Thorndykes Gesetz nicht mehr viel zu tun. Vor allem gilt es nicht in dieser Allgemeinheit (vgl. Westmeyer, 1973) bzw. ist es schlichtweg Unsinn: Dann müsste nämlich z.B. ein vollgefressenes Versuchstier, das in der Skinner-Box mit Elektroschocks gequält wird, lernen, ständig jenen Hebel zu betätigen, durch den es noch weiter Futterpillen bekommt, statt andere Verhaltensweisen auszuprobieren, bis es endlich jene findet, mit der es dem Elektroschock entgehen kann.

Eine Auflösung der Zirkularität des Gesetzes des Effektes wird allerdings möglich, wenn wir den Behaviorismus verlassen:

1. Mit einer Stimulusveränderung ist das Hinzufügen oder Wegnehmen eines Stimulus gemeint

Dazu brauchen wir weder physiologische Messnormen der Triebstärke, noch die (wenig sinnvolle) Annahme, dass Verstärker transsituational sind, sondern wir müssen auf die *Semantik* der Stimuli und Verhaltensweisen zurückgreifen:

Jenseits des Behaviorismus: die Aufhebung der Zirkularität

Thorndykes Gesetz des Effektes kann dann so umformuliert werden,

- dass Verknüpfungen von Situationen und Reaktionen, die von einer (im Darwin'schen Sinne) funktionalen Situationsveränderung gefolgt sind, verstärkt werden,
- während die Stärke von Verknüpfungen, die von einer (im Darwin'schen Sinne) disfunktionalen Situationsveränderung gefolgt sind, abnimmt.

Thorndykes satisfying state of affairs haben wir darin durch den Terminus ‚funktionale Situationsveränderung‘ ersetzt, den annoying state of affairs durch den Terminus ‚disfunktionale Situationsveränderung‘.

Was eine funktionale oder disfunktionale Situationsveränderung ist, wird dabei auf der Grundlage allgemeiner empirischer Gesetzmäßigkeiten zu den beobachtungssprachlich beschriebenen Stimuli und Verhaltensweisen hinzukonstruiert:

- Funktional sind z.B. all jene Situationsveränderungen, durch welche der Organismus Zugang zu fitnessbegrenzenden Ressourcen erhält,
- disfunktional sind dagegen die Situationsveränderungen, durch welche ihm der Zugang zu fitnessbegrenzenden Ressourcen vorenthalten wird.

Welche Ressourcen als fitnessbegrenzende Ressourcen zu gelten haben, ist aus unserem biologischen (Gesetzes-)Wissen ableitbar, dass das Überleben des Individuums und/oder die Arterhaltung nachhaltig gefährdet sind, wenn dem Organismus (im Singular und/oder Plural) diese Ressourcen dauerhaft vorenthalten werden.

Das entspricht genau den Versuchsanordnungen, derer sich auch der Behaviorismus schon immer bedient hat: Nahrungsdeprivierte Organismen werden Reaktionen erlernen, durch welche sie sich Zugang zu Futter verschaffen können. Sexuell deprivierte Organismen werden Reaktionen erlernen, die ihnen Zugang zu einem Sexualpartner verschaffen usw.

3.7.2 Zirkularität, Pseudoempirie und strukturelle Theorieanteile

Der Zirkularitätsverdacht gegenüber der Frustrations-Aggressions-Hypothese

Könnte die Zirkularität des Gesetzes des Effektes geheilt werden, indem wir die Semantik der Stimuli und Verhaltensweisen berücksichtigt haben, so entsteht der Zirkularitätsverdacht in anderen theoretischen Kontexten damit erst.

Betrachten wir z.B. die Frustrations-Aggressions-Hypothese in ihrer ursprünglichen Fassung, so ist unmittelbar ersichtlich, dass zwischen Frustration und Aggression im Sinne von Dollard et al. (1939) kein struktureller Zusammenhang besteht:

- Unter Frustration verstehen die Autoren die Störung einer zielgerichteten Aktivität.
- Unter Aggression verstehen sie eine Verhaltenssequenz, die auf die Verletzung eines Organismus oder Organismusersatzes abzielt.

Die Terminologie der Yale-Gruppe stellt also durchaus eine geeignete Basis dar, um die Frustrations-Aggressions-Hypothese als empirische Hypothese behaupten zu können.

Berücksichtigen wir dagegen die Semantik der Stimuli und Verhaltensweisen und sehen das wesentliche Bestimmungsmerkmal der Frustration in der Vorenthaltung fitnessbegrenzender Ressourcen, während Aggression als eine Form des Konkurrenzverhaltens um eben solche Ressourcen zu verstehen ist, so ist die Unabhängigkeit der beiden Sachverhalte zumindest fraglich.

In der Humanpsychologie, wo wir es nicht nur mit funktionalem Verhalten, sondern mit zielgerichtetem Handeln und folglich mit individuellen Sinngehalten zu tun haben, verschärft sich der Verdacht noch, weshalb z.B. Smedslund (1976) in seinem Vortrag auf dem Internationalen Kongress für Psychologie in Paris die Vermutung geäußert hat, dass die Relation zwischen Frustration und Aggression bereits a priori bestehe, also *strukturell* sei und keinerlei empirischen Gehalt habe.

Die Suche nach dem strukturellen Kern der Theorie

Dass Frustration und Aggression keine unabhängigen Sachverhalte darstellen, bedeutet nicht notwendigerweise, dass die Theorie zirkulär ist; wohl jedoch, dass sie (zumindest) einen strukturellen Kern besitzt, den es erst noch im Detail herauszuarbeiten gilt.

Wie wir bereits am Beispiel von Sieglers Theorie der kognitiven Entwicklung im Kindesalter gesehen haben, muss die Existenz eines strukturellen Theoriekerns keineswegs von Nachteil für den empirischen Gehalt der Theorie sein. Strukturelles und empirisches Wissen können einander nicht ersetzen, und die Identifikation struktureller Theorieanteile kann empirische Forschung nicht obsolet machen, sondern im Gegenteil: Sie ist ein wichtiges Instrument zur Vermeidung von Pseudoempirie und zur Steigerung der Effizienz und Präzision der empirischen Forschung. In diesem Sinne werden wir im Anschluss an Kempf (1978, 1996) zeigen, dass die Frustrations-Aggressions-Hypothese tatsächlich einen strukturellen Kern besitzt, dass dieser strukturelle Kern die Formulierung weit präziserer empirischer Hypothesen erlaubt, und dass er zugleich die Integration der vielen verschiedenen Aggressionshypothesen, die es in der Literatur gibt, in eine umfassende und geschlossene Aggressionstheorie ermöglicht.

3.7.3 Frustration und Aggression

Wenn es eine A-priori-Relation zwischen Frustration und Aggression gibt, dann muss diese bereits mit den Mitteln der Logik und der Terminologie (sowie falls benötigt: der Mathematik) beweisbar sein. Um dies überprüfen zu können, bedürfen wir zu allererst einer geklärten Terminologie. Zu diesem Zwecke greifen wir auf die folgenden Definitionen zurück:

Definitionen

D1: Ein Konflikt (lat. *confligere* = zusammenstoßen, streiten) besteht im Aufeinandertreffen miteinander unvereinbarer Handlungstendenzen (Handlungen oder Ziele).

Bestehen die Handlungstendenzen innerhalb einer Person, so spricht man von einem inneren oder intrapersonalen Konflikt. Die Unverträglichkeit der Handlungen oder Ziele zweier oder mehrerer Akteure (Personen, Gruppen oder Institutionen) bezeichnet man dagegen als sozialen Konflikt. Soziale Konflikte können kooperativ oder kompetitiv ausgeglichen werden.

Eine Kooperationssituation besteht, wenn die Zielerreichung der Akteure positiv korreliert ist. Im Falle der reinen Kooperation kann jeder Teilnehmer sein Ziel dann und nur dann erreichen, wenn die anderen, mit denen er verbunden ist, ihre Ziele ebenfalls erreichen (win-win-Situation).

Eine Konkurrenzsituation besteht, wenn die Zielerreichung der Akteure negativ korreliert ist. Im Falle der reinen Konkurrenz kann jeder Teilnehmer sein Ziel dann und nur dann erreichen, wenn die anderen, mit denen er verbunden ist, ihre Ziele nicht erreichen können (win-lose-Situation).

- D2: Unter einer Frustration (lat. *frustror* = vereiteln) verstehen wir ein Ereignis, als dessen Wirkung eine Handlung verhindert wird und/oder in Hinblick auf die mit ihr verfolgten Ziele erfolglos bleibt.
- D3: Eine Aggression (lat. *aggredior*) besteht in der (versuchten) Durchsetzung eigener Handlungen oder Ziele gegen den (tatsächlichen oder vermeintlichen) Willen eines anderen.
- D4: Gewalt (engl. *violence*; lat. *violo* = Gewalt antun, wehe tun, misshandeln, verletzen; beflecken, besudeln; verletzen, versehren, schänden, entweihen, beleidigen) besteht in der Verletzung der körperlichen oder seelischen Unversehrtheit eines anderen.

Der strukturelle Kern der Frustrations-Aggressions-Hypothese

Aufgrund dieser Definitionen ergeben sich die folgenden strukturellen Gesetzmäßigkeiten:

- G1: *Frustration und Konflikt*. Tritt eine Frustration (tatsächlich oder vermeintlich) infolge des Handelns eines anderen ein und deutet der Akteur diese Folge als vom anderen intendiert, so besteht nach Meinung des Akteurs ein Konflikt (D1, D2).
- G2: *Konflikt und Aggression*. Beharrt der Akteur in diesem (tatsächlichen oder vermeintlichen) Konflikt weiterhin auf seinen ursprünglichen Handlungszielen, so ist jede Handlung, welche auf die unmittelbare Durchsetzung dieser Ziele gerichtet ist, per definitionem eine Aggression (D1, D3).
- G3: *Katharsis*. Setzt der Akteur seine Handlungsziele damit erfolgreich durch, dann ist die Handlung abgeschlossen, und es bedarf (zur Erreichung dieser Ziele) keiner weiteren Handlungen, also auch keiner weiteren Aggressionen (D1).

Der Zusammenhang zwischen Frustration und Aggression ist an bestimmte subjektseitige Prämissen gebunden; konkret:

- an die Deutung der Frustration als vom anderen intendiert, und

- an die (versuchte) Durchsetzung der ursprünglichen eigenen Ziele.
- G4: *Aggression und Aggression*. Die Durchsetzung ursprünglicher Handlungsziele in Frustrationssituationen ist per se noch keine Aggression (D3). Zur Aggression wird sie erst, wenn der Akteur meint, sie gegen den Willen eines anderen (der die Erreichung dieses Zieles verhindern will) durchzusetzen (D3).
Ist dies der Fall, so ist diese Aggression vom Standpunkt des Akteurs aus die Verteidigung seiner ursprünglichen Handlungsziele gegen eine (tatsächliche oder vermeintliche) Aggression des anderen, mit welcher dieser die Zielerreichung verhindert hat (D3).
- Die Aggressions-
Aggressions-
Hypothese
- G5: *Attribution und Aggression*. Wenn jemand andere für seine Misserfolge (Frustrationen) verantwortlich macht und ihnen dabei böse Absichten unterstellt, wird er die Frustrationen aggressiv beantworten (sofern er nicht auf sein ursprüngliches Handlungsziel verzichtet) (G1, G2, G4).
- Die Aggressions-
Attributions-
Hypothese
- G6: *Bedrohung und Aggression*. Der strukturelle Zusammenhang zwischen Frustration und Aggression besteht nicht nur für bereits eingetretene Frustrationen, sondern auch für drohende Frustrationen, denn entscheidend dafür, dass die Durchsetzung eigener Handlungsziele eine Aggression darstellt, ist
- nicht, ob der Akteure in seiner Zielerreichung bereits behindert worden ist,
 - sondern dass der Akteur meint, seine Ziele gegen den Willen eines Konfliktpartners durchsetzen zu müssen (D3).
- Letzteres wiederum folgt daraus, ob sich der Akteur in einer Konkurrenzsituation (win-lose) oder in einer Kooperations-situation (win-win) zu befinden meint (D1).
- G7: *Verschiebung der Aggression*. Insbesondere in Konkurrenz-situationen zwischen mehr als zwei Akteuren besteht nur selten eine reine Konkurrenz.
- Verschiebung der
Aggression
- Wird ein Akteur (A) an seiner Zielerreichung durch einen anderen Akteur (B) gehindert (Frustration), und
 - meint er, seine Ziele aber dennoch gegen einen Dritten

3. Soft Science: Die Herausforderung an die Sozialwissenschaften

(C) durchsetzen zu können, und

- beharrt er auf der Durchsetzung seiner Ziele,
- dann muss sich diese Durchsetzung nicht gegen B richten,
- sondern es kann zur Verschiebung der Aggression gegen C kommen.

Diese kann auch dann eintreten, wenn A gar nicht der Meinung ist, dass B die Frustration intendiert hat.

Empirische
Prämissen

Mit den oben rekonstruierten strukturellen Gesetzmäßigkeiten haben wir die innere Logik von Konflikten bereits ein Stück weit analysiert, ohne dass wir dazu irgendwelche Beobachtungen gebraucht haben. Insofern handelt es sich dabei um aristotelische Wissensanteile: Wir haben zuerst die wesentlichen Bestimmungsmerkmale von Konflikt, Frustration, Aggression und Gewalt herausgearbeitet (D1-D4) und dann daraus die strukturellen Gesetzmäßigkeiten (G1-G6) (sach-)logisch hergeleitet.

Aggression und
Gewalt

Dabei haben wir gesehen, dass diese Gesetzmäßigkeiten für die (subjektseitig definierte) Aggression gelten. Wie man sich leicht überzeugen kann, gelten sie jedoch nicht für die (objektseitig definierte) Gewalt.

Darin zeigt sich die Zweckmäßigkeit der vorher erarbeiteten terminologischen Unterscheidung zwischen Aggression und Gewalt.

Dem Behaviorismus, der diese Unterscheidung nicht treffen konnte, weil er nur objektseitig definierte Beobachtungstermini als zulässig erachtete und Aggression daher von ihren schädigenden Folgen (Gewalt) her zu definieren versuchte, war es nicht möglich, diese strukturellen Gesetzmäßigkeiten zu erkennen.

Im Behaviorismus und in allen empirischen Untersuchungen, die Aggression (als AV) über die schädigenden Folgen operationalisieren, gelten diese Gesetzmäßigkeiten folglich nicht mehr strukturell, sondern nur noch statistisch, und zwar:

- proportional zu der Wahrscheinlichkeit, mit der Aggression (in der jeweiligen Versuchsanordnung) zu Gewalt eskaliert, sowie
- proportional zu der Wahrscheinlichkeit, mit der die in den strukturellen Gesetzmäßigkeiten enthaltenen subjektseitigen Prämissen durch die jeweilige Versuchsanordnung (Manipulation der UV) provoziert werden konnten.

Dass die empirischen Ergebnisse dieser Experimente teils widersprüchlich sind, stellt keine Falsifikation der strukturellen Gesetzmäßigkeiten dar, sondern demonstriert lediglich die Mangelhaftigkeit der in den Experimenten verwendeten Operationalisierungen.

Mit der Analyse der Logik von Konflikten wird empirische Forschung dennoch *nicht obsolet*, sondern versetzt uns im Gegenteil in die Lage, weit präzisere und intelligenter empirische Hypothesen zu formulieren, als dies in der Variablenpsychologie der Fall ist, die meint, der Theorienbildung sei schon genüge getan, wenn man einen statistischen Zusammenhang zwischen (mehr oder minder) aus der Luft gegriffenen UVn und AVn postuliert.

In den oben erarbeiteten strukturellen Gesetzmäßigkeiten ist eine Reihe von subjektseitigen Prämissen enthalten:

Subjektseitige
Prämissen

1. Frustration führt dann zu Aggression (G1-G2),
 - *wenn* die Frustration als vom andern intendiert gedeutet wird, und
 - *wenn* der Akteur an den frustrierten Handlungszielen festhält.
2. Zur Verteidigung gegen eine Aggression des anderen (G4-G5) wird die Aggression dann,
 - *wenn* der andere für die Frustration verantwortlich gemacht wird, und
 - *wenn* ihm dabei böse Absicht unterstellt wird.
3. Bedrohung wird mit Aggression beantwortet (G6),
 - *wenn* der Akteur meint, sich in einer Konkurrenzsituation zu befinden.
4. Zur Verschiebung der Aggression (G7) kann es kommen,
 - *wenn* A auf der Durchsetzung seiner Ziele beharrt, und
 - *wenn* er meint, diese gegen einen Dritten durchsetzen zu können.
5. *Ob* es zur Verschiebung der Aggression kommt, wird wesentlich davon abhängen,
 - welche Erfolgchancen er für die Durchsetzung seiner Handlungsziele sieht: sei es gegen C, gegen B oder sei es auf einem kooperativen Weg, der sich gegen keinen der

beiden wendet, und

- welche (positiven oder negativen) Nebenwirkungen er erwartet.

Empirische
Fragestellungen

Weil es sich bei diesen subjektseitigen Prämissen aber um nichts anderes als um die empirischen Voraussetzungen handelt, unter welchen die strukturellen Gesetzmäßigkeiten zum Tragen kommen, ergibt sich daraus unmittelbar eine Vielzahl von weiterführenden empirischen Fragestellungen, wie z.B.:

1. Unter welchen Bedingungen neigen Vpn dazu, Frustration als von einem anderen intendiert zu deuten (d.h. andere für die Frustration verantwortlich zu machen und ihnen dabei böse Absichten zu unterstellen)?
2. Unter welchen Bedingungen neigen Vpn dazu, an den frustrierten (oder bedrohten) Handlungszielen festzuhalten (d.h. auf der Durchsetzung ihrer Ziele zu beharren)?
3. Unter welchen Bedingungen neigen Vpn dazu, Konflikte als Konkurrenzsituation zu deuten (in der sie sich durchsetzen müssen)?
4. Unter welchen Bedingungen eskaliert die Aggression zur Gewalt?
5. Unter welchen Bedingungen neigen Vpn zu der Meinung, ihre Ziele (nur) mit gewaltförmigen Mitteln durchsetzen zu können?
6. Welche Bedingungen sind für die Beurteilung der Erfolgchancen maßgeblich?
7. Welche (positiven oder negativen) Nebenwirkungen begünstigen die Wahl gewaltförmiger Mittel?

3.7.4 Exkurs: Die Lerntheorie der Aggression

Einige dieser Bedingungen wurden von der Arbeitsgruppe um Albert Bandura (1979) im Rahmen seiner Lerntheorie der Aggression untersucht. Diese Theorie geht davon aus, dass die gewaltförmige Eskalation aggressiver Interaktionen durch eine Vielzahl an kulturellen Faktoren begünstigt wird. Dazu gehören soziale Normen, welche die Durchsetzung um jeden Preis vorschreiben, ebenso wie soziale Modelle (Vorbilder), welche Kompetenzen einer kooperativen Konfliktbearbeitung vernachlässigen und/oder die (vermeintliche) Funktionalität von Gewalt als Mittel der Konfliktlösung vor Augen führen.

Bandura und Walters (1959) verglichen die Erziehungspraktiken der Eltern von Heranwachsenden, die zu Gewalt neigten, mit denen der Eltern von nicht auffälligen Jugendlichen. Beide Gruppen von Jugendlichen stammten aus intakten Mittelschichtelternhäusern, in denen die Verhaltensstandards der Umgebung gesetzestreuere Verhalten förderten. Verglichen wurden die Familien von Jugendlichen, die wiederholt antisoziales Aggressionsverhalten gezeigt hatten, mit denen von Jugendlichen, die weder gewaltbereit noch passiv waren.

Soziale Normen

- Die Eltern der nicht-gewaltbereiten Jugendlichen ermutigten ihre Söhne durchaus, ihre Prinzipien entschlossen zu verteidigen, aber sie entschuldigten gewaltförmige Aggression als Mittel zur Beilegung von Streitigkeiten nicht und bekräftigten ihre Söhne selten, wenn sie als Reaktion auf Provokation von gewaltförmigen Aggressionen Gebrauch machten.
- Die Eltern der zu Gewalt neigenden Jugendlichen tolerierten zwar zu Hause keinerlei gewaltförmige Handlungsweisen, entschuldigten jedoch provokative und gewaltförmige Handlungen gegenüber anderen, spornten ihre Kinder aktiv dazu an und belohnten sie dafür. Während der frühen Kindheit nahm die Anstiftung häufig sogar die Form von Forderungen an, nach denen der Junge seine Fäuste gebrauchen sollte, wenn er sich mit seinen Gegnern abgab.

Ähnliche Ergebnisse erbrachte eine Untersuchung der Erziehungspraktiken in den Familien gewaltbereiter und gehemmter kleiner Jungen (Bandura, 1960).

Solche Normen werden nicht nur durch Erziehung, sondern auch durch die Medien und innerhalb der Peer-Group weitergegeben. In der Sozialisation Heranwachsender gehen soziale Bekräftigung und Modellierungseinflüsse oft Hand in Hand.

So zeigte sich in der Studie von Bandura und Walters ebenfalls, dass sich die Familien der zu Gewalt neigenden und der nicht-gewaltbereiten Jugendlichen in ihrem Vorbildcharakter deutlich voneinander unterschieden:

Soziale Modelle

- In den familiären Interaktionen modellierten die Eltern der nicht-gewaltbereiten Jugendlichen rücksichtsvolles Verhalten und vertrauten hauptsächlich auf Argumente, wenn sie sich mit sozialen Problemen beschäftigten.
- Im Unterschied dazu zeigten die Eltern der gewaltbereiten

Jungen zwar wenig antisoziales Aggressionsverhalten, aber sie modellierten und verstärkten immer wieder kämpferische Verhaltensweisen.

Kontrollfragen

- Wie lautet Thorndykes Formulierung des Gesetzes des Effektes?
- Welche Voraussetzungen müssen u.a. erfüllt sein, damit es sich dabei tatsächlich um eine Erklärung nach dem HO-Schema handelt?
- Welchen Vorwurf hat Postman gegenüber Thorndykes Gesetz erhoben?
- Inwiefern liegt dieser Vorwurf etwas schief?
- Inwiefern ist die Zulässigkeit von Thorndykes Formulierung des Gesetzes für den Behaviorismus zweifelhaft?
- Nennen Sie zwei Versuche, das Gesetz des Effektes streng behavioristisch zu formulieren, und begründen Sie, warum diese Versuche gescheitert sind!
- Wie kann man die Zirkularität von Skinners Gesetz des operanten Konditionierens aufheben? Was ist davon zu halten?
- Wie kann man das Gesetz des Effektes jenseits des Behaviorismus zirkelfrei formulieren?
- Wie können Thorndykes Konzepte des satisfying bzw. annoying state of affairs verhaltensbiologisch definiert werden?
- Diskutieren Sie, ob die Frustrations-Aggressions-Hypothese in der Formulierung von Dollard et al. (1939) als empirische Hypothese behauptet werden kann!
- Erläutern Sie, warum die (sachlogische) Unabhängigkeit von Frustration und Aggression bei Berücksichtigung der Semantik der Stimuli und Verhaltensweisen fraglich wird!
- Erläutern Sie den Unterschied zwischen intrapersonalen und sozialen Konflikten!
- Erläutern Sie den Unterschied zwischen Kooperations- und Konkurrenzsituationen!
- Stellen Sie den strukturellen Zusammenhang zwischen
 - a) Frustration und Konflikt,
 - b) Konflikt und Aggression,
 - c) Aggression und Katharsis,
 - d) Aggression und Aggression,
 - e) Attribution und Aggression,
 - f) Bedrohung und Aggression
 dar und erläutern Sie,
 - a) auf welchen terminologischen Voraussetzungen er beruht, und
 - b) an welche subjektseitigen Prämissen er ggf. gebunden ist!
- Erläutern Sie, was man unter der Verschiebung der Aggression versteht und unter welchen subjektseitigen Prämissen es dazu kommen kann!
- Erläutern Sie anhand eines Beispiels, warum strukturelle Gesetzmäßigkeiten immer nur relativ zu einer bestimmten Terminologie bewie-

sen werden können!

- Welche Aussagekraft haben Experimente bzw. empirische Untersuchungen, die Aggression über die schädigenden Folgen operationalisieren, bezüglich der Geltung der strukturellen Gesetzmäßigkeiten, welche zwischen Frustration und Aggression bestehen?
- Begründen Sie, warum empirische Forschung mit der Analyse der Logik von Konflikten nicht obsolet wird!
- Nennen Sie einige empirische Fragestellungen, die sich unmittelbar aus der Analyse der Logik von Konflikten ergeben!
- Nennen Sie zwei der kulturellen Faktoren, durch welche die gewaltförmige Eskalation aggressiver Interaktionen nach Bandura begünstigt wird!
- Skizzieren Sie die Ergebnisse der Untersuchung von Bandura und Walters!

Handlungstheoretische Prozessanalyse

3.8.1 Die Erfahrungsgebundenheit struktureller Theorieanteile

Strukturelle Theorieanteile können per definitionem nicht an der Erfahrung scheitern. Gleichwohl sind sie zumindest insofern erfahrungsgebunden, als sie nicht einfach am Schreibtisch konstruiert werden können. Schon die Entscheidung darüber, welche terminologischen Unterscheidungen wir treffen, misst sich an der Zweckmäßigkeit dieser Unterscheidungen und ist damit auf Erfahrungen angewiesen.

Als es darum ging, die Frustrations-Aggressions-Hypothese strukturell zu rekonstruieren, setzten wir uns ja auch nicht einfach an den Schreibtisch und überlegten gleichsam aus dem hohlen Bauch heraus, womit Aggression in welcher Weise zusammenhängen könnte, sondern wir hatten schon vorher (zumindest) die (Alltags-)Erfahrung gemacht, dass Frustration häufig Aggressionen heraufbeschwört. Im Falle der Frustrations-Aggressions-Hypothese verfügten wir sogar über systematische empirische und experimentelle Erfahrungen; und das, was wir aufgrund dieser Erfahrungen bereits vermuteten, versuchten wir nun zu erklären, indem wir die strukturellen von den empirischen Anteilen trennten, um darüber die empirischen Bedingungen genauer herausarbeiten (und prüfen) zu können, unter denen Frustration tatsächlich zu Aggression führt.

Damit wissen wir nun nicht nur, dass Frustration häufig zu Aggression (und nicht selten sogar zur Verletzung eines anderen) führt, sondern wir haben uns ein strukturelles Wissen um die Logik von Konflikten und ein empirisches Wissen um einige der Bedingungen, unter denen diese Logik zum Tragen kommt, erarbeitet.

Auch im weiteren Forschungsprozess werden wir uns nicht mit der (statistisch-induktiven) Erklärung zufrieden geben, dass z.B. unter bestimmten Sozialisationsbedingungen eher mit antisozialem

Aggressionsverhalten zu rechnen ist als unter anderen, sondern wir werden eine Erklärung dafür anstreben, *warum* unter diesen Sozialisationsbedingungen die gewaltförmige Eskalation von Konflikten vorprogrammiert ist. Dabei werden wir einerseits wieder auf strukturelle Faktoren stoßen, welche der Eskalationsdynamik von Konflikten inhärent sind (Eskalationslogik); andererseits werden wir typische Begründungsmuster finden, durch welche die Gewaltanwendung (subjektiv) legitimiert wird (intentionale Erklärung), und wir werden drittens typische Interpretationsprozesse rekonstruieren, wie die objektseitig definierten Umweltbedingungen sich zu Situationen konstituieren, in denen diese Begründungsmuster handlungsrelevant werden (narrative Erklärung).

3.8.2 Exkurs: Die Dynamik der Konflikteskalation

Als Beispiel betrachten wir die Dynamik der Konflikteskalation, die sich nach Glasl (1992) in neun Stufen vollzieht, die in drei Hauptebenen (Creighton, 1992) gegliedert werden können (vgl. Abb. 3.8.1).

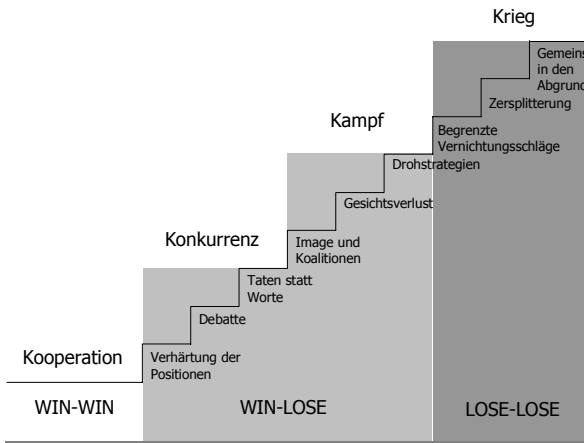


Abb. 3.8.1: Dynamik der Konflikteskalation nach Glasl (1992) und Creighton (1992)

Der typische Eskalationsverlauf beginnt damit, dass der Konflikt von einer oder beiden Parteien als Konkurrenzsituation nach dem win-lose-Prinzip konzeptualisiert wird. Es kommt zum Streit. Das Ziel der gemeinsamen Problemlösung wird durch die Absicht der Durchsetzung eigener Positionen verdrängt. Es wird wichtiger, den Streit zu gewinnen, als das Problem zu lösen. Die Konflikt-

parteien empfinden einander als Gegner, nicht mehr als Team. Das vorherrschende Klima ist das eines Wettkampfes, den man nur gewinnen oder verlieren kann.

Glasl unterscheidet die Stufen:

Verhärtung: Noch überwiegt die Kooperation, doch die Standpunkte prallen zuweilen aufeinander.

Debatte: Denken, Fühlen und Wollen der Konfliktparteien polarisieren sich. Zwischen kooperativen und kompetitiven Einstellungen besteht ein labiles Gleichgewicht.

Taten statt Worte: Das Konkurrenzverhalten gewinnt die Oberhand. Eine Strategie der vollendeten Tatsachen wird eingeschlagen.

Kampf

Wenn der Punkt erreicht ist, an dem miteinander reden nicht mehr hilft, sondern Fakten geschaffen werden müssen, dann ist das Ende einer konstruktiven Konfliktbearbeitung erreicht, die sich um Sachfragen dreht und die die Parteien allein bewältigen können. Von nun an wird die Beziehung zur gegnerischen Partei zentraler Konfliktgegenstand (Müller & Büttner, 1996). Der nun stattfindende Übergang zum Kampf erfolgt, wenn eine der Parteien sich verletzt fühlt und befürchtet, noch mehr verletzt zu werden. Ziel der Auseinandersetzung wird es, die andere ebenfalls zu verletzen. Keine der Parteien ist bereit, eigene Schwächen einzugehen. Das Bild der gegnerischen Partei wird verzerrt. Standen im Streit noch Sachfragen im Vordergrund, so werden jetzt vor allem die Einstellungen der Konfliktparteien zum Problem.

Glasl unterscheidet die Stufen:

Sorge um Image und Koalition: Die Konfliktparteien drängen einander in negative Rollen. Sie werben um Anhänger und Koalitionen.

Gesichtsverlust: Die ganze Person des Gegners erscheint – auch rückwirkend – in einem neuen Licht.

Drohstrategien: Gewaltdenken und Gewalthandeln nehmen zu. Die Parteien versuchen, durch Drohungen aufeinander Einfluss auszuüben.

Krieg

Der Kampf eskaliert zum Krieg, wenn die Absicht zur physischen oder psychischen Vernichtung des Gegners entsteht und/oder dieser unter Anwendung von Gewalt zum Nachgeben gezwungen

werden soll. Die Parteien sehen sich nicht mehr als Menschen, sondern als Feinde. Jedes gesprochene Wort (selbst Schweigen) wird zur Waffe, jedes gehörte Wort (auch Schweigen) wird als Waffeneinsatz interpretiert. Sogar Bemühungen einer Partei, das Ausmaß des Konfliktes zu dämpfen, werden nur noch als taktische Manöver oder Manipulationsversuche wahrgenommen. Am Ende geht es nicht einmal mehr darum, auf Kosten des anderen zu gewinnen, sondern nur noch darum, dass der andere nicht gewinnen darf (lose-lose-Prinzip). Wird diese Schwelle überschritten, dann haben die Parteien voneinander keine menschlichen Bilder mehr, sondern sehen sich als Objekte, die einander nur noch loswerden wollen. Von nun an wird die gegenseitige Gewaltanwendung zum vorherrschenden Konfliktgegenstand (Müller & Büttner, 1996).

Glasl beschreibt diesen Prozess in den Stufen:

Begrenzte Vernichtungsschläge: Die Werte verkehren sich ins Gegenteil, relativ geringerer eigener Schaden wird bereits als Gewinn angesehen.

Zersplitterung: Paralyisierung und Desintegration des feindlichen Systems werden zum vorherrschenden Ziel.

Gemeinsam in den Abgrund: Es führt kein Weg mehr zurück. Die Konfliktparteien betreiben die Vernichtung des Feindes selbst um den Preis der Selbstvernichtung.

3.8.3 Narrative Erklärung, Eskalationslogik und Begründungsmuster

Ähnlich wie die kognitive Entwicklung durch Sieglers Informationsverarbeitungsmodelle (Kapitel 3.6.3) wird die Eskalationsdynamik von Konflikten durch die Eskalationsmodelle von Creighton und Glasl (Kapitel 3.8.2) dargestellt, indem die Geschichte ihres (typischen) Verlaufes erzählt wird.

Typische
Verlaufsgeschichten

Auf den ersten Blick mag es so scheinen, dass diese typischen Verlaufsgeschichten durch *Generalisierung* zustande kommen und also erzählen, wie die kognitive Entwicklung bzw. die Eskalation von Konflikten *im Allgemeinen* abläuft.

Tatsächlich entstehen solche typischen Verlaufsgeschichten aber nicht durch den Prozess der Generalisierung, sondern durch den

Prozess der *Abstraktion*, d.h. indem von den Besonderheiten des Einzelfalles abgesehen wird.

Nonlineare
Konfliktverläufe

So ist der Prozess der Konflikteskalation zumeist kein linearer Prozess, der Stufe für Stufe voranschreitet, sondern er vollzieht sich eher nach dem Muster "zwei Schritte vor, einen Schritt zurück". Auf jeder Stufe der Konflikteskalation besteht grundsätzlich die Möglichkeit zur Umkehr, zur schrittweisen Deeskalation des Konfliktes. Dies ist allerdings umso schwieriger, je weiter die Eskalation fortgeschritten ist und in hoch eskalierten Konflikten ohne die Vermittlung einer neutralen dritten Partei kaum noch möglich.

Insbesondere dann, wenn nur eine der Konfliktparteien versucht, den Konflikt ein Stück weit zu deeskalieren, sind Enttäuschungen vorprogrammiert. Indem das Eskalationsniveau immer noch hoch ist, erscheint die Frustration der Deeskalationsbemühungen als neuerlicher Beweis für die Verächtlichkeit des Gegners und bestätigt gleichzeitig die Gerechtigkeit der eigenen Anliegen. Die Eskalationsdynamik kommt erneut und mit doppelter Wucht in Gang.

Die Frage nach dem
Typischen

Dies wirft die Frage auf, wie man unter diesen Voraussetzungen überhaupt von einem typischen Eskalationsverlauf sprechen kann bzw. was das Prädikat ‚typisch‘ hier eigentlich bedeutet. Die Antwort darauf finden wir, indem wir einen Blick darauf werfen, worin das Typische der von Glasl und Creighton beschriebenen Eskalationsverläufe besteht und die empirischen von den strukturellen Anteilen zu trennen versuchen.

- *Der empirische Gehalt der Eskalationsmodelle* besteht darin, dass es die beschriebenen Stufen (Glasl) bzw. Ebenen (Creighton) tatsächlich gibt und dass sie in eskalierenden Konflikten mehr oder weniger regelmäßig anzutreffen sind.
- *Der strukturelle Gehalt der Eskalationsmodelle* besteht darin, dass den beschriebenen Stufen und Ebenen eine gewisse Eskalationslogik zu eigen ist.

Die Eskalationslogik

Eine Ebene folgt auf die andere, indem die Parteien immer drastischere Mittel zur Durchsetzung ihrer Ziele anwenden. Je drastischer die Mittel sind, desto mehr bedürfen sie einer Rechtfertigung, welche die eigenen Ziele als gleichermaßen gerecht wie unverzichtbar erscheinen lässt und die eigene Seite in der Rolle des Opfers sieht, den Gegner dagegen in der des Angreifers, gegen welchen es sich zu verteidigen gilt und der nichts zu

befürchten hätte, wenn er nicht an ungerechtfertigten Zielen festhalten wollte und seine Handlungen nicht so verabscheuenswürdig wären. Doch je verächtlicher der Gegner erscheint, desto weniger gemeinsame Anknüpfungspunkte gibt es und desto drastischeres Mittel bedarf es, um sich gegen seine Angriffe zu verteidigen usw.

Auf einen einfachen Nenner gebracht, besagt die Eskalationslogik nichts anderes, als dass der Konflikt auf der nächst höheren Stufe bereits weiter eskaliert ist als auf der jeweils vorangegangenen. D.h. sie definiert, was wir überhaupt unter der Eskalation von Konflikten verstehen.

Diese Eskalationslogik, die an anderer Stelle (Kempf, 2003) im Detail herausgearbeitet ist, bedeutet jedoch *weder*, dass Konflikte stets eskalieren müssen, *noch* bedeutet sie, dass der Eskalationsprozess alle diese Ebenen – eine nach der anderen – durchlaufen muss. Sie bedeutet noch nicht einmal, dass der Konflikt als solcher durch einen bestimmten Eskalationsgrad beschreibbar ist.

Asynchrone
Konfliktverläufe

Aufgrund der gestörten Kommunikation zwischen den Konfliktparteien verläuft die Eskalation oft nicht auf beiden Seiten synchron. Bewegen sich die Parteien auf unterschiedlichem Eskalationsniveau, so heizt dies die Eskalationsdynamik noch weiter an: Die Partei, die sich auf der höheren Eskalationsstufe bewegt, interpretiert die Handlungen ihres Gegners im Sinne ihres eigenen Eskalationsniveaus und unterstellt ihm böse Absichten, welche weit darüber hinausgehen, was er tatsächlich im Sinn hat. Der anderen Partei, die sich auf dem niedrigeren Eskalationsniveau bewegt, erscheinen die Handlungen ihres Gegners dagegen als überzogen und unangemessen und bestätigen damit, dass die Verächtlichkeit des Gegners noch weit größer ist als befürchtet.

Die dargestellten Eskalationsmodelle erzählen also die Geschichte des Konfliktverlaufes in *idealisierter* Form, d.h. unter der (gedachten) Voraussetzung, dass der Konflikt tatsächlich eskaliert, dass die Eskalation des Konfliktes auf beiden Seiten synchron abläuft, dass die verschiedenen Ebenen der Konflikteskalation eine nach der anderen durchlaufen werden usw.

Unvermittelte
Gewaltausbrüche

Sind diese Voraussetzungen im konkreten Fall erfüllt und die Konflikteskalation verläuft so wie in den Eskalationsmodellen beschrieben, dann reicht der Verweis auf die Eskalationsmodelle zu ihrer Erklärung völlig aus. Ist dies jedoch nicht der Fall, so ent-

steht ein Erklärungsbedarf dafür, dass der konkrete Konfliktverlauf vom typischen Eskalationsverlauf abweicht.

So sind z.B. asynchrone Konfliktverläufe häufig die Folge von Machtungleichgewicht zwischen den Konfliktparteien. Während die mächtigere Partei ihre Rechte und Ziele durchsetzen kann, ohne den Konflikt wirklich austragen zu müssen, ist die machtärmere Partei ständigen Frustrationen ausgesetzt und befindet sich in einem andauernden Zustand der Deprivation, aus welchem sich mit maßvollen Mittel zu befreien zunehmend aussichtslos erscheint, bis es am Ende zu scheinbar unvermittelten Gewaltausbrüchen kommt.

Im Extremfall kann es sogar sein, dass die mächtigere Partei das Bestehen eines Konfliktes überhaupt nicht bemerkt hat und sich in einem Zustand vollständiger Harmonie mit dem Partner zu befinden meint, bis sie der Konflikt plötzlich und mit völlig unverständlicher Schärfe trifft.

Narrative Erklärung

Die Erklärung des abweichenden Konfliktverlaufes erfolgt nun ebenfalls wieder durch das Erzählen einer Geschichte, d.h.

- soweit sie den konkreten Einzelfall betrifft, durch das Erzählen einer konkreten Geschichte, und
- soweit wir von Besonderheiten des Einzelfalles abstrahieren können, durch das Erzählen einer typischen Geschichte.

Machtungleichgewicht

Da Machtlosigkeit auch bedeutet, dass man mit Reden wenig oder nichts erreichen kann, führt Machtlosigkeit nur allzu oft dazu, dass man die Konflikte gegenüber dem Partner gar nicht mehr anspricht. Die Fixierung auf die eigenen Rechte und Ziele bis hin zu deren Idealisierung und die Verurteilung der gegnerischen Handlungen (und seien sie auch noch so kooperativ gemeint) bis hin zu deren Dämonisierung und zur Bestreitung der Rechte des Gegners vollziehen sich im Verborgenen und werden allenfalls gegenüber dritten Parteien thematisiert, von denen man sich Unterstützung erhofft. Indem diese Unterstützung nicht selten gewährt wird, statt vermittelnd in den Konflikt einzugreifen, finden die sich herausbildenden Vorurteile gegenüber dem Gegner soziale Unterstützung und werden langsam zur Gewissheit. Wenn der Konflikt schließlich ausbricht, geschieht dies mit solcher Vehemenz, dass die konkreten Sachfragen, welche den (ursprünglichen) Konflikthalt ausmachen, hinter dem Ausmaß der Feindseligkeit zurücktreten, ja mitunter nicht einmal mehr konkret be-

nannt werden können – oder allenfalls als lange Liste von Beweisen für die Verabscheuenswürdigkeit des Gegners. Aber die Frustrationen, welche die Beziehung belasten, liegen in der Vergangenheit, und darum geht es jetzt nicht mehr. Die Beziehung selbst steht zur Disposition, der frühere Partner ist zum Feind geworden.

Mit dem Erzählen dieser Geschichte stellen wir eine sinnstiftende Verbindung her zwischen den beiden Eckpunkten der vermeintlichen Harmonie, in der sich die mächtigere Partei mit ihrem Partner zu befinden glaubt, und der unverhohlenen Feindseligkeit, mit der dieser ihr plötzlich begegnet. Begründungsmuster

Wenn wir im Kontext der narrativen Erklärung z.B. darauf verweisen, dass der machtärmere Partner den Konflikt so lange unangesprochen lässt, "weil man mit Reden eh nichts erreichen kann, wenn der andere um so vieles mächtiger ist", so berufen wir uns damit auf ein mehr oder minder verbreitetes Begründungsmuster, mit dem wir zu rechnen haben; das wir im abstrakten Fall als idealisierende Voraussetzung unterstellen können, dessen tatsächliche Befolgung wir im konkreten Fall jedoch erst handlungstheoretisch rekonstruieren müssen.

Um die Konfliktodynamik zu erklären, welche sich daraus ergibt, genügt es aber nicht, nur die Sichtweise der einen oder der anderen Konfliktpartei nachzuvollziehen. Die Geschichten 1. und 2. Person, welche das subjektive Erleben der Konfliktparteien darstellen, liefern nur das Rohmaterial für die Geschichte 3. Person, welche die Interaktion der Konfliktparteien erklärt.

Kontrollfragen

- Diskutieren Sie anhand eines Beispiels die Erfahrungsgebundenheit struktureller Theorieanteile!
- Skizzieren Sie die Dynamik der Konflikteskalation nach Glasl und Creighton!
- Was bedeutet Perspektivendivergenz der Akteure?
- Welchen strukturellen Grund hat diese Perspektivendivergenz?
- Erläutern Sie am Beispiel des Eskalationsmodells von Glasl und Creighton, inwieweit solche Prozessmodelle
 - a) durch Generalisierung und/oder
 - b) durch Abstraktion zustandekommen!
- Erläutern Sie, worin
 - a) der empirische und

b) der strukturelle Gehalt
der Eskalationsmodelle besteht!

- Worin besteht die Eskalationslogik, die den Modellen zugrundeliegt, und was besagt sie?
- Erläutern Sie die logische Struktur der Erklärung abweichender Konfliktverläufe!
- Erläutern Sie, in welcher Weise Machtungleichgewicht zu asynchronen Konfliktverläufen führen kann und wie sich dies auf die Eskalationsdynamik des Konfliktes auswirkt!

3.9

Qualitative Forschungsmethoden

Sobald es darum geht, typische Eskalationsverläufe herauszuarbeiten und/oder aktuelle Abweichungen davon zu rekonstruieren, sehen wir uns mit einer Situation konfrontiert, in der wir es nicht mehr mit isolierten Handlungen in einer (experimentell) kontrollierten Handlungsumgebung, sondern mit komplexen Handlungs- und Lebenszusammenhängen zu tun haben, deren Erklärung es erforderlich macht, die subjektive Bedeutung zu rekonstruieren, welche diese Handlungszusammenhänge und die in ihnen ausgeführten Handlungen für die Probanden selbst haben.

Rekonstruktion von
Bedeutungen

In methodologischer Hinsicht macht dies ein radikales Umdenken erforderlich.

Schon die Reflexionstermini, mit denen wir z.B. Werte und Normen beschreiben, die an der Konstitution der subjektseitig bestehenden Situation und damit an der Sinnstrukturierung der Handlungszusammenhänge beteiligt sind,

- sind einer terminologischen Festschreibung nicht zugänglich (vgl. Kapitel 3.3.1), und
- was jemand meint, wenn er solche Worte verwendet, was die Worte eigentlich bedeuten, kann nur durch eine methodisch kontrollierte Interpretation erhellt werden.

Gerade jene Aspekte der subjektiven Welt des Menschen, die wir in in der experimentellen Forschung als Störvariablen auszuschalten versucht haben, treten in ihrer Naturwüchsigkeit jetzt ins Zentrum des Forschungsinteresses.

Die daraus resultierende Notwendigkeit eines methodologischen Paradigmenwechsels verschärft sich noch, wenn wir – im Grenzbereich zur Soziologie – die gesellschaftliche Konstruktion der Wirklichkeit zum Forschungsgegenstand machen und nach den sozialen bzw. kulturellen Regeln fragen, nach denen sie sich vollzieht.

Entsprechend ist es auch nicht weiter verwunderlich, dass es zuerst die Kulturanthropologie, die Ethnologie und die Soziologie waren, die das Programm einer interpretativen Sozialforschung ausgearbeitet haben, während sogenannte qualitative Methoden in der Psychologie bis in die 70er Jahre des vergangenen Jahrhunderts hinein allenfalls in den Bereich der explorativen Forschung verwiesen wurden, wo ihnen ein gewisser heuristischer Wert für die Hypothesengenerierung zugesprochen wurde.

Kulturabhängigkeit

Kulturabhängigkeit – etwa des Eskalationsverlaufes von Konflikten – muss nicht notwendigerweise bedeuten, dass bestimmte Verhaltensweisen in bestimmten Kulturen gänzlich fehlen oder durch andere ersetzt sind. Kulturabhängigkeit kann sich gegebenenfalls auch darin manifestieren, dass die Verhaltensweisen in anderer Reihenfolge aufeinanderfolgen und das entsprechende Verhalten dadurch eine andere Bedeutung erhält. Schon dies allein kann interkulturelle Missverständnisse erzeugen und/oder Konflikte hervorbringen.

Ein anschauliches Beispiel hierfür geben Watzlawick et al. (1967): Unter den während des II. Weltkrieges in England stationierten US-amerikanischen Soldaten war die Ansicht verbreitet, dass die Engländerinnen sexuell überaus leicht zugänglich seien. Gleichzeitig behaupteten die Engländerinnen ihrerseits, dass die US-Soldaten übertrieben stürmisch seien.

Eine Untersuchung, an der u.a. Margaret Mead mitarbeitete, führte zu einer interessanten Lösung des Widerspruchs: Wie sich herausstellte, durchläuft das Paarungsverhalten – vom ersten Kennenlernen der Partner bis zum Geschlechtsverkehr – in England wie in den USA ungefähr dreißig verschiedene Verhaltensformen, wobei aber die Reihenfolge dieser Verhaltensformen in den beiden Kulturen verschieden ist. Während z.B. das Küssen in den USA bereits sehr früh kommt (etwa auf Stufe 5), tritt es im britischen Paarungsverhalten erst sehr spät auf (etwa auf Stufe 25). Wenn also ein US-Soldat eine Engländerin zu küssen versuchte, so war dies für ihn noch ein ziemlich harmloser Annäherungsversuch, während er andererseits der Engländerin, die sich dadurch bereits in höchstem Maße sexuell bedrängt fühlte, als sehr stürmisch erschien. Die Engländerin selbst war damit praktisch vor die Entscheidung gestellt, ob sie die Beziehung an diesem Punkt abbrechen oder sich dem Partner sexuell hingeben sollte. Entschied sie sich für die letztere Alternative, so fand sich

der US-Soldat mit einem Verhalten konfrontiert, das für ihn durchaus nicht in dieses Frühstadium der Beziehung passte und das ihm geradezu als schamlos erscheinen musste.

3.9.1 Interpretativität

Die Bedeutung eines Sachverhaltes konstituiert sich durch *Interpretation*, d.h. dadurch, dass dieser Sachverhalt in einen bestimmten *Kontext* gestellt wird. Damit verweist jede Bedeutung reflexiv auf das Ganze, und die Bedeutung eines Handelns oder eines sprachlichen Ausdrucks wird nach Lamnek (1988, 25) "nur durch den Rekurs auf den (symbolischen oder sozialen) Kontext seiner Erscheinung verständlich".

Der kommunikative Grundcharakter der Sozialforschung

Um die (subjektseitige) Bedeutung eines Sachverhaltes (für einen Probanden) zu verstehen, ist es daher erforderlich, die vom Probanden selbst vollzogenen Interpretationsleistungen nachzuvollziehen, d.h. jene Kontexte zu rekonstruieren, in denen der jeweilige Sachverhalt aus der Sicht des Probanden zu sehen ist.

Schütze (1978, 118) spricht deshalb vom kommunikativen Grundcharakter der Sozialforschung: "Der kommunikative Sozialforscher behandelt das informierende Gesellschaftsmitglied [den Probanden, W.K.] als prinzipiell orientierungs-, deutungs- und theoriemächtiges Subjekt". Mit seinen Interpretationsleistungen deutet der Proband aber nicht nur die ihm zugängliche Wirklichkeit, sondern konstituiert sie zugleich: Die Sicht der Wirklichkeit ist perspektivenabhängig, und mit dem Wechsel der Perspektive ändert sich auch das, was als wirklich gilt.

Daraus ergibt sich eine Reihe von methodologischen Prinzipien:

Methodologische Prinzipien

1. *Das Prinzip der Kommunikation:* Wer die bedeutungsstrukturierte Welt eines anderen erfassen will, muss mit dem anderen kommunizieren und dabei dafür Sorge tragen, dass er in der ihm vertrauten Sprache und nach den ihm vertrauten Regeln kommunikativer Verständigung sprechen kann.
2. *Das Prinzip der Fremdheit:* Wer die bedeutungsstrukturierte Welt eines anderen erfassen will, tut gut daran, diese in ihrer jeweiligen Eigenart, potentiellen Andersheit und Fremdheit zu betrachten und ihr nicht unmittelbar die eigenen Begriffe, Kategorien und Interpretationsmuster überzustülpen.
3. *Das Prinzip der Offenheit:* Wer über die bedeutungsstrukturierte

rierte Welt eines anderen etwas erfahren will, muss seine Forschungspraxis für die Selbstthematierungen des anderen weitgehend offen halten.

4. *Prinzip der Reflexion*: Wer seine Forschungspraxis für die Selbstthematierungen des anderen offen halten will, muss die eigenen (oft impliziten) Vorannahmen ständig reflektieren, offen legen und in Frage stellen.

Methodische
Konsequenzen

Diese Prinzipien finden sowohl in den Erhebungs- als auch in den Analysemethoden der qualitativen Sozialforschung ihren Niederschlag und führen dort zu einer Vorgehensweise, welche jener der quantitativen Sozialforschung z.T. diametral entgegengesetzt ist.

Dies beginnt bereits bei der Stichprobenauswahl mit der Bevorzugung der Typikalität gegenüber der Repräsentativität und des Theoretical Sampling (Glaser & Strauss, 1967) gegenüber einer Zufallsauswahl der Probanden.

Es setzt sich fort mit der Bevorzugung offener Erhebungsinstrumente (Witzel, 1982) wie der teilnehmenden Beobachtung (vgl. einführend Lamnek, 1989; Lüders, 2000) und des Rollenspiels als dessen experimenteller Variante (z.B. Volmerg et al., 1983), des offenen Interviews in seinen verschiedenen Formen (vgl. einführend Lamnek, 1989; Hopf, 2000), insbesondere des narrativen Interviews (Schütze, 1977), und – unter Einbeziehung von Techniken der klientenzentrierten Gesprächsführung nach Rogers (1972) – des partnerzentrierten Gesprächs (Kempf, 1987) sowie der Gruppendiskussion (vgl. einführend Volmerg, 1977; Bohnsack, 2000). Und es gipfelt schließlich in einer flexiblen Handhabung interpretativer Analysemethoden, für die es – anders als z.B. in der statistischen Datenanalyse – kein festes Schema gibt. "Die Praxis qualitativer Forschung ist generell dadurch geprägt, dass es nicht *die* Methode gibt, sondern ein methodisches Spektrum unterschiedlicher Ansätze, die je nach Fragestellung und Forschungstradition ausgewählt" (Flick et al., 2000, 22) und gegenstandsangemessen modifiziert, variiert und abgewandelt werden können, so dass sich qualitative Forschung als besonders resistent gegenüber allen Versuchen einer Reduzierung auf starre methodische Anleitungen und Regeln erweist.

Die (angebliche)
Explorationsfunktion
der qualitativen
Sozialforschung

Die Notwendigkeit einer flexiblen und gegenstandsangemessenen Handhabung methodischer Ansätze betrifft insbesondere auch das Prinzip der Offenheit, das bis Ende der 80er Jahre des vergangenen Jahrhunderts vorrangig durch den "Verzicht auf

eine Hypothesenbildung ex ante" (Hoffmann-Riem, 1980, 343ff) operationalisiert und mit der (angeblichen) Explorationsfunktion qualitativer Sozialforschung begründet wurde.

So spricht Hoffmann-Riem (a.a.O., 345) davon, dass Sozialforschung für den interpretativ arbeitenden Soziologen "weitgehend Exploration" sei, und Lamnek (1988, 23) zieht daraus die Konsequenz, dass sich "qualitative Sozialforschung ... nicht als hypothesenprüfendes, sondern als hypothesengenerierendes Verfahren [versteht]. Die Hypothesen sollen ... erst auf der Grundlage der im Forschungsprozess erhobenen Daten entwickelt und nicht wie in der quantitativen Sozialforschung üblich, die vorabformulierten Hypothesen auf der Grundlage der Daten empirisch überprüft werden".

Im Unterschied dazu hatte mit Siegfried Kracauer (1952) bereits einer der Begründer der qualitativen Sozialforschung die Forderung erhoben, den qualitativen Forschungsmethoden ihren eigenen Rang und Status zuzuerkennen, und Meinefeld (2000) kommt zu dem Schluss, dass die Entscheidung gegen Ex-ante-Hypothesen zwar zu einer Konsolidierung der qualitativen Position in Abgrenzung von der quantitativen Methodologie beigetragen, aber zugleich auch zu einer erkenntnistheoretisch nicht haltbaren Festlegung geführt habe, welche die Einsatzmöglichkeiten qualitativer Sozialforschung unangemessen begrenzte.

Tatsächlich werden in der Gegenüberstellung von Offenheit vs. Ex-ante-Hypothesen verschiedene Dinge durcheinander gebracht.

- Dazu gehört erstens, dass das Prinzip der Offenheit für die explorative Felderkundung per definitionem unverzichtbar ist, für die (experimentelle) Hypothesenprüfung dagegen bestenfalls irrelevant.
- Zweitens stellt die Hypothesenformulierung ex ante zwar eine substantielle Voraussetzung des Prüfexperimentes dar, für die Felderkundung ist sie jedoch verzichtbar.
- Drittens können die Forderung nach Ex-ante-Hypothesen und die Forderung nach Offenheit des Forschungsprozesses zwar in Widerspruch zueinander geraten, schließen einander jedoch nicht grundsätzlich aus. Zumindest aber können offene mit geschlossenen Erhebungsinstrumenten, offene mit geschlossenen Analyseschritten und hypothesengenerierende mit hypothesentestenden Fragestellungen kombiniert werden.

Der (angebliche) Antagonismus von Offenheit und Ex-ante-Hypothesen

- Viertens haben wir es in den Sozialwissenschaften mit einer Vielzahl von verschiedenen Arten von Hypothesen zu tun, die sich bereits in ihrer sprachlichen und logischen Struktur voneinander unterscheiden und entsprechend auch verschiedene Schritte der Geltungsprüfung bzw. Geltungsbegründung erfordern.
- Fünftens ist die Unterscheidung zwischen hypothesenprüfenden und hypothesengenerierenden Verfahren insofern irreführend, als sie die den Eindruck erweckt, es würde sich dabei um dieselbe Art von Hypothesen handeln.
- Sechstens entspricht dieses Selbstverständnis auch nicht der tatsächlichen Vorgehensweise der interpretativen Sozialforschung, die damit in den Bereich der bloßen Voruntersuchungen verwiesen wird.
- Siebentens verhindert es jegliche Akkumulation von Wissen innerhalb des interpretativen Paradigmas, so dass jede qualitative Untersuchung wieder bei Null beginnen müsste, und
- nicht zuletzt nährt dieses Selbstverständnis die Illusion, dass sich die empirische Wirklichkeit dem interpretativen Sozialforscher gleichsam von selbst offenbaren würde.

Die Illusion der Voraussetzungslosigkeit

Ganz im Sinne dieser Illusion schreibt z.B. Lamnek (1989, 194f.) mit Bezug auf die qualitative Inhaltsanalyse, dass die theoretischen Konzepte und Hypothesen dort "nicht aufgrund von wissenschaftlichem und alltagsweltlichem Vorwissen formuliert, sondern durch kontrolliertes Fremdverstehen der von den Untersuchten verwendeten Alltagskonzepte generiert" würden. Die qualitative Inhaltsanalyse versuche derart, "den Inhalt selbst sprechen zu lassen und aus ihm heraus die Analyse zu entfalten".

Dabei wird übersehen, dass jedes Fremdverstehen zumindest ein alltagsweltliches Vorwissen voraussetzt. D.h. die Interpretation setzt voraus, dass der Interpret einen Vorbegriff der Sache hat, und zwar in dreierlei Hinsicht (vgl. Mollenhauer & Rittelmeyer, 1977, 59):

- als *Verstehbarkeit des semantischen Gehalts* (der Bedeutung) der verbalen (und ggf. nonverbalen) Zeichen (Worte, Sätze, Gesten) der (Umgangs-)Sprache, die der Interpret mit dem Text teilt,
- als *Klassifikation des Textes* (autobiographische Erzählung, Rollenspielsituation, Gruppendiskussion etc.), der verstanden werden soll und so als ein Gegenstand mit einer bestimmten,

alltagspraktisch eingespielten gesellschaftlichen Bedeutung eingeordnet wird,

- als *theoretisch reflektierter Vorbegriff*, der eine aus der wissenschaftlichen Arbeit gewonnene Annahme darüber enthält, welche Art von Hypothesen am Text entwickelt und/oder überprüft werden kann.

Methodisch kontrolliertes Fremdverstehen setzt darüber hinaus voraus, dass dieses Vorwissen die Interpretationen nicht hinter dem Rücken des Interpreten beeinflussen und/oder verzerren darf. Dies ist aber nur möglich, wenn der Interpret sein Vorwissen systematisch reflektiert, in Frage stellt und ggf. modifiziert. Gerade das ist aber ausgeschlossen, wenn man sich der Illusion hingibt, man könne ohne Vorwissen an den Text herangehen und den Inhalt für sich selbst sprechen lassen.

3.9.2 Text und Textanalyse

Interpretative Sozialforschung ist Textwissenschaft im weitesten Sinne (auch die Daten aus der teilnehmenden Beobachtung liegen uns in Form von Texten, nämlich Beobachtungsprotokollen, vor).

Der Begriff des Textes ist in der Psychologie und in den Sozialwissenschaften im Allgemeinen sehr weit gefasst: Er verweist nicht nur auf Sprache als Transportmittel von Text, sondern auf Sprache als *Ort* der Sinnkonstitution.

Konkret bedeutet dies,

"... dass ‚meaning making‘ (Bruner) nicht in einem Individuum passiert, welches dem Wissenschaftler sodann darüber Auskunft gibt, vielmehr wird vor unseren Ohren und Augen in der Interaktion gemeinsam Sinn konstituiert" (Kraus, 2001, 242).

Alltagssprachlich wird der Begriff des Textes zumeist auf verschriftlichte Sprache angewendet. Dies ist jedoch ein sehr enger Textbegriff, der z.B. beim Interview nur auf die Transkription des gesprochenen Wortes anwendbar ist. Eine erste Erweiterung stellt die Einbeziehung der parasprachlichen Kommunikation dar, z.B. Lachen, Hüsteln, Schweigen (durch Audioaufnahmen noch zugängliche Lautäußerungen, die durch Anwendung besonderer Transkriptionsregeln mitprotokolliert werden können). Eine nächste Erweiterung ist die Einbeziehung der nichtsprachlichen (non-verbalen) Kommunikation, z.B. Mimik, Gestik (Videoaufnahmen

erforderlich). Noch grundsätzlicher wird der Textbegriff durch die Semiologie erweitert, die letztlich jede Art von Zeichensystemen unter den Textbegriff fasst. Insofern können auch bildliche Darstellungen und Dinge als Zeichen bzw. als Zeichensysteme gelesen werden.

Für die Analyse von Texten sind in den Sprach- und Sozialwissenschaften eine Vielzahl von Methoden entwickelt worden. Diese Methoden beruhen z.T. auf unterschiedlichen theoretischen (disziplinären) Bezugsrahmen und Erkenntnisinteressen. So gehen Sprachwissenschaftler an die Analyse von Texten anders heran als Soziologen, und die wiederum anders als Psychologen, wobei es sich allerdings nicht um konkurrierende Herangehensweisen handelt, sondern um Herangehensweisen, die – je nach ihrem disziplinären Kontext – verschiedene Facetten des Textes fokussieren und sich wechselseitig ergänzen. Zum Teil spielen auch methodologische Interessen eine Rolle. So definiert z.B. Berelson (1954, deutsch nach Lamnek, 1989, 180):

"Inhaltsanalyse ist eine Forschungstechnik zur objektiven, systematischen und quantitativen Deskription von manifestem Kommunikationsinhalt."

Gemeinsam ist all diesen Methoden, dass sie relativ zu ihren jeweiligen theoretischen und methodologischen Grundannahmen begründet sind und daher, wie Kraus (2001) betont, nicht unabhängig von diesen angewendet werden können. Gleichwohl gibt es eine Reihe von Leitgedanken, zu denen sich jede textanalytische Methode so oder so verhalten muss.

Der hermeneutische
Zirkel

Der Interpretationsvorgang verläuft nach den Regeln einer auf Verständigung zielenden Kommunikation und enthält zwei aufeinander bezogene Operationen:

- Einerseits werden aus der Perspektive des Interpreten *Fragen* an den Text gerichtet,
- andererseits steckt in der *Antwort* immer auch mehr oder anderes, als gefragt wurde.

Folglich ergeben sich aus dem Text gleichsam *Rückfragen* an die Angemessenheit, Ergiebigkeit etc. der vom Interpreten an den Text herangetragenen Gesichtspunkte. Ein angemessenes Textverständnis ist daher nicht auf Anhieb möglich, sondern erst im Ergebnis einer Interpretationsarbeit, "in der in vielen einzelnen Schritten immer wieder der Versuch unternommen werden muss, Sinn und Bedeutung des ‚Ganzen‘ aus seinen einzelnen Teilen

und die Bedeutungsstellung dieser im Rahmen des Ganzen zu ermitteln" (Mollenhauer & Rittelmeyer, 1977, 59). Diese in der Natur des interpersonalen Verstehens begründete wechselseitige Erschließung von Teil und Ganzem wird in der Literatur als hermeneutischer Zirkel bezeichnet:

Die methodische Vorgehensweise der Textinterpretation erfordert einen ständigen Wechsel vom Detail zum Ganzen und zurück. Die einzelne Textstelle erschließt sich in ihrer Bedeutung erst, wenn der Gesamtkontext bekannt ist, und die Bedeutung des Gesamtkontextes erschließt sich erst aus der Bedeutung der einzelnen Textstellen. Am Anfang haben die Interpretationen noch einen ziemlich hypothetischen bis spekulativen Charakter. Indem sie Schritt für Schritt überprüft, präzisiert und ggf. modifiziert werden, steht schließlich am Ende des Interpretationsvorganges das, was vom gesamten Text eingelöst werden kann.

Wegen der Kontextabhängigkeit von Bedeutungen sind Verfahrensweisen methodisch unzulässig, die lediglich einzelne Textstellen herausgreifen und zur Illustration von Hypothesen oder Theorien verwenden (Heinze & Klusemann, 1980).

Die Rekonstruktion von Kontexten

Die möglichen Interpretationen eines Textes werden durch

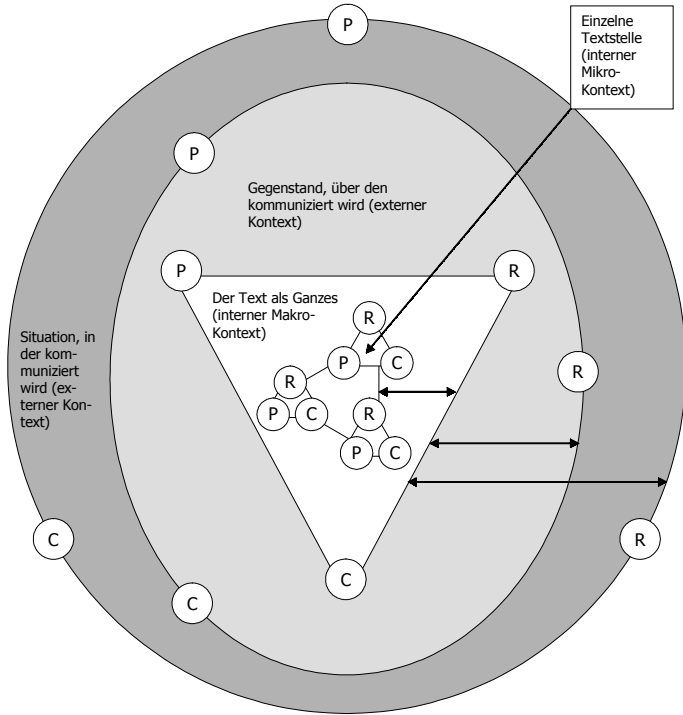
- den Text selbst (interner Kontext),
- den (verfügbaren) externen Kontext, auf den der Text verweist, und durch
- die interpretationsleitenden Fragen begrenzt, welche der Interpret an den Text richtet.

Bezugnehmend auf Bühler und Watzlawick et al. (vgl. Kapitel 3.4.5) können wir

- bezüglich des internen Kontextes zwischen der einzelnen Textstelle (interner Mikro-Kontext) und dem Text als Ganzem (interner Makro-Kontext) sowie
- bezüglich des externen Kontextes zwischen der Situation, in der kommuniziert wird, und dem Gegenstand, über den kommuniziert wird,

unterscheiden. Dabei ist jeder dieser Kontexte sowohl unter dem Aspekt seiner Darstellungsfunktion (report aspect), unter dem Aspekt seiner Ausdrucksfunktion (command aspect) und unter dem Aspekt seiner Appellfunktion (parade aspect) zu berücksichtigen (vgl. Abb. 3.9.1).

Abb. 3.9.1: Interner und externer Kontext
 R = report aspect,
 C = command aspect,
 P = parade aspect



Weil sich diese Kontexte und ihre verschiedenen Aspekte wechselseitig interpretieren, erfordert der oben dargestellte hermeneutische Zirkel

- nicht nur einen ständigen Wechsel zwischen der einzelnen Textstelle und dem Text als Ganzem,
- sondern auch einen ständigen Wechsel zwischen dem Text und seinen externen Kontexten
- sowie einen ständigen Wechsel zwischen der Fokussierung der Darstellungs-, Ausdrucks- und Appellfunktion sowohl des Textes als auch seiner Kontexte und sowohl der einzelnen Textstelle als auch des Textes als Ganzem.

Konsistenz als methodisches Prinzip

Die Annahme einer gewissen Konsistenz der Bedeutungen ist nach Heinze & Klusemann (1980, 103) keine empirische Frage, sondern "ein fundamentales Theorem des erkennenden Umgangs mit Dingen überhaupt", d.h. ein methodisches Prinzip, unter dem die gültigen Interpretationen eines Textes innerhalb der möglichen Interpretationen ausgezeichnet werden können.

- So kann die Übereinstimmung zwischen Textstellen als Indiz dafür genommen werden, dass sie richtig verstanden wurden.
- Andererseits bilden gerade die Widersprüchlichkeiten und Ungereimtheiten den Anknüpfungspunkt für eine weitergehende Analyse des Textes.

Widersprüche dürfen daher nicht einfach geglättet werden, sondern zunächst müssen alle Bedeutungsgehalte, die eine Textstelle haben kann, aufgelistet und als gleichwertig betrachtet werden.

Alltagspraktische Verständnisbildung verlässt sich auf das Maß konventionell eingespielter Deutungen und tendiert dazu, mögliche Missverständnisse unkenntlich zu machen, indem das gegenseitige Verstehen als unhinterfragtes Axiom vorausgesetzt wird. Die qualitative Sozialforschung versucht, diese internalisierte Intersubjektivität (Leithäuser & Volmerg, 1977) durch systematisches Misstrauen

Durchbrechen
internalisierter
Intersubjektivität

- gegenüber den sich anbietenden Interpretationen,
- gegenüber den eigenen Voreinstellungen, Erwartungen etc. sowie
- gegenüber der Vollständigkeit des erfassten Kontextwissens

zu durchbrechen, wobei

- nicht nur Sinn und Bedeutung des Textes expliziert, sondern auch
- deren soziale Genese und ideologisch-gesellschaftliche Funktion rekonstruiert werden (Heinze & Klusemann, 1980).

Da "wir nie sicher sein können, die gültige Interpretation gefunden zu haben, da wir nicht ausschließen können, dass es weitere Kontext-Informationen geben könnte, die uns zu einer Änderung unserer Deutung veranlassen müssten" (a.a.O., 104), kann das Ergebnis einer Interpretation immer nur vorläufig sein und hat somit "den wissenschaftslogischen Status einer Hypothese. Unter keinen Umständen kann mit zweifelsfreier Gewissheit und hinreichenden Gründen behauptet werden, dass eine vorgelegte Deutung richtig ist" (Mollenhauer & Rittelmeyer, 1977, 59).

Hypothesenförmigkeit
der Ergebnisse

Die Untersuchungsergebnisse qualitativer Studien unterscheiden sich diesbezüglich nicht von den Ergebnissen quantitativer Untersuchungen, und auch die methodologischen Schlussfolgerungen, welche die qualitative Sozialforschung daraus zieht, sind dieselben, namentlich "dass die Interpretation (wie jede andere wis-

senschaftliche Behauptung auch) wenigstens so vorgenommen und vorgetragen werden muss, dass sie falsifizierbar ist, das heißt, dass unzutreffende Interpretationen erkannt und verworfen werden können" (a.a.O., 59f). Das ist nur möglich, wenn das Interpretationsverfahren durchsichtig, die einzelnen Interpretationsschritte nachvollziehbar und die verwendeten Termini in ihrer Bedeutung bestimmt und eindeutig auf die einzelnen Textbestandteile beziehbar sind.

Kontrollfragen

- Diskutieren Sie, warum die Rekonstruktion von Konfliktverläufen ein radikales Umdenken in methodologischer Hinsicht erforderlich macht!
- Erläutern Sie, wie sich dabei das Forschungsinteresse ändert!
- Inwiefern ist damit zu rechnen, dass die Eskalationsstufen nach Glas kulturabhängig sein könnten, und worin könnte diese Kulturabhängigkeit zum Ausdruck kommen?
- Erläutern Sie anhand eines Beispiels, wie abweichende kulturelle Regeln zu interkulturellen Missverständnissen und Konflikten führen können!
- Was versteht man in der qualitativen Sozialforschung unter Interpretativität? Wie ist sie begründet, und welche methodischen Prinzipien ergeben sich daraus?
- Was versteht man in der qualitativen Sozialforschung unter den Prinzipien
 - a) der Kommunikation,
 - b) der Fremdheit,
 - c) der Offenheit und
 - c) der Reflexion?
- Wie schlagen sich die o.g. Prinzipien auf
 - a) die Stichprobenauswahl,
 - b) die Wahl der Erhebungsinstrumente und
 - c) die interpretativen Analyseverfahren nieder?
- Diskutieren Sie den vermeintlichen Antagonismus von Offenheit und Ex-ante-Hypothesen!
- Diskutieren Sie, inwiefern die Interpretation voraussetzt, dass der Interpret bereits einen Vorbegriff der Sache hat!
- Kritisieren Sie die Auffassung, wonach qualitative Sozialforschung ihre theoretischen Konzepte und Hypothesen nicht aufgrund von wissenschaftlichen und/oder alltagsweltlichem Vorwissen formuliere, sondern versuche, "den Inhalt selbst sprechen zu lassen und aus ihm heraus die Analyse zu entfalten"!
- Erläutern Sie den Begriff des Textes!
- Erläutern Sie den Unterschied zwischen verbaler, parasprachlicher und nichtsprachlicher Kommunikation!
- Worin unterscheiden sich die verschiedenen in den Sprach- und Sozi-

alwissenschaften entwickelten Methoden der Textanalyse? Was ist ihnen gemeinsam, und was folgt daraus?

- Erläutern Sie, was man unter einem hermeneutischen Zirkel versteht und wie dieser bei Berücksichtigung der metakommunikativen Axiomen von Watzlawick et al. zu formulieren ist!
- Welche Rolle spielt die Konsistenz der Bedeutungen für die Textinterpretation?
- Was bedeutet der Terminus ‚internalisierte Intersubjektivität‘?
- Erläutern Sie, warum das Ergebnis einer Interpretation den wissenschaftslogischen Status einer Hypothese hat, und diskutieren Sie, welche Konsequenzen sich daraus ergeben!

Literatur

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317-332.
- Andersen, E.B. (1973). *Conditional inference and models for measuring*. Copenhagen: Mentalhygiejnisk Forlag.
- Andersen, E.B. (1982). Latent structure analysis – A survey. *Scandinavian Journal of Statistics*, 9, 1-12.
- Andrich, D. (1978a). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, 2, 581-594.
- Andrich, D. (1978b). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Appelsmeyer, H., Kochinka, A., Straub, J. (2000). *Qualitative Methoden*. In: Straub, J., Kempf, W., Werbik, H. (Hg.). *Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven*. 3. Auflage. München: Deutscher Taschenbuchverlag, 709-742.
- Asch, S.E. (1951). Effects of group pressure on the modification and distortion of judgements. In: Guetzkow, H. (ed.). *Groups, leadership and men*. Pittsburgh: Carnegie.
- Asch, S.E. (1956). Studies of independence and conformity: A minority of one against an unanimous majority. *Psychological Monographs*, 70/9 (whole no. 416).
- Aschenbach, G. (1985). Teleologische und methodologische Überlegungen zur Psychoanalyse – Zu einer handlungstheoretischen Rekonstruktion. In: Aschenbach, G., Kempf, W. (Hg.). *Methodenprobleme in der Psychotherapieforschung*. Eschborn: Fachbuchhandlung für Psychologie, 131-206.
- Balogh, A. (1994). *Ansätze zur Verbesserung der Diagnostik bei Problemlöseprozessen am Beispiel von Balkenwaageaufgaben unter Verwendung der Latent-Class-Analyse*. Universität Konstanz: Diplomarbeit.
- Bandura, A. (1960). Relationship of family patterns to child behavior disorders. Progress Report, Stanford University, Project No. M-1734, United States Public Health Service.
- Bandura, A. (1979). *Aggression. Eine sozial-lerntheoretische Analyse*. Stuttgart: Klett-Cotta.
- Bandura, A. Walters, R.H. (1959). *Adolescent aggression*. New York: Ronald Press.

- Bergius, R. (1965). Behavioristische Motivationsbegriffe. In: Thomae, H. (Hg.). Handbuch der Psychologie in 12 Bänden. Band 2, S. 817-866. Göttingen: Hogrefe.
- Berelson, B. (1954). Content analysis. In: Lindzey, G. (ed.). Handbook of social psychology. Vol. 1. Theory and method. Cambridge, Mass.: Addison-Wesley, 488-522.
- Berkowitz, L. (1969). Roots of Aggression. New York: Atherton.
- Binet, A., Simon, Th. (1905). Methodes nouvelles pour le diagnostique du niveau intellectuel des anormaux. *Année Psychologique*, 11, 191-244.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In: Lord, F.M., Novick, M.R. (1968). Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley. Part 5, 397-479.
- Bischof, N. (1981). Aristoteles, Galilei, Kurt Lewin – und die Folgen. In: Michaelis, W. (Hg.). Bericht über den 32. Kongress der Deutschen Gesellschaft für Psychologie in Zürich 1980. Band 1. Göttingen: Hogrefe, 17-39.
- Bischof, N. (1991). Gescheiter als all die Laffen. Ein Psychogramm von Konrad Lorenz. Hamburg: Rasch & Röhring.
- Blumer, H. (1973). Der methodologische Standpunkt des symbolischen Interaktionismus. In: Arbeitsgruppe Bielefelder Soziologen (Hg.). Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit. Band 1. Symbolischer Interaktionismus und Ethnomethodologie. Reinbek: Rowohlt, 80-146.
- Bohnsack, R. (2000). Gruppendiskussion. In: Flick, U., Kardorff, E.v., Steinke, I. (Hg.). Qualitative Sozialforschung. Ein Handbuch. Reinbek: Rowohlt, 369-383.
- Bortz, J., Döring, N. (1995). Forschungsmethoden und Evaluation für Sozialwissenschaftler. 2. Auflage. Berlin: Springer.
- Brandstädter, J. (1982). Apriorische Elemente in psychologischen Forschungsprogrammen. *Zeitschrift für Sozialpsychologie*, 13, 267-277.
- Brandstädter, J. (1984). Apriorische Elemente in psychologischen Forschungsprogrammen: Weiterführende Argumente und Beispiele. *Zeitschrift für Sozialpsychologie*, 15, 151-158.
- Brandstädter, J. (1987). "A rose has no teeth" – Zum Problem der Unterscheidung zwischen Begriffsverwirrung und überraschenden empirischen Befunden in der Psychologie. In: Brandstädter, J. (Hg.). Struktur und Erfahrung in der psychologischen Forschung. Berlin: de Gruyter.
- Brandstädter, J., Greve, W. (1999). Intentionale und nichtintentionale Aspekte des Handelns. In: Straub, J., Werbik, H. (Hg.). Handlungstheorie. Begriff und Erklärung des Handelns im interdisziplinären Diskurs. Frankfurt/Main: Campus, 185-212.
- Bühler, K. (1934). Sprachtheorie. Stuttgart: Fischer.

- Buss, A.H. (1961). *The psychology of aggression*. New York: Wiley & Sons.
- Campbell, D.T., Stanley, J.C. (1963). Experimental and quasi-experimental designs for research in teaching. In: Gage, N.L. (ed.). *Handbook of research in teaching*. Chicago: Rand McNally, 171-246.
- Carrier, M., Mittelstraß, J. (1989). *Geist, Gehirn, Verhalten. Das Leib-Seele-Problem und die Philosophie der Psychologie*. Berlin: de Gruyter.
- Coleman, S.R. (1981). Historical context and systematic functions of the concept of the operant. *Behaviorism*, 9, 207-226.
- Cramér, H. (1946). *Mathematical methods in statistics*. Princeton: Princeton University Press.
- Creighton, J.L. (1992). *Schlag nicht die Türe zu. Konflikte aushalten lernen*. Reinbek: Rowohlt.
- Crott, H. (1979). *Soziale Interaktion und Gruppenprozesse*. Berlin: Kohlhammer.
- Danto, A.C. (1974). *Analytische Philosophie der Geschichte*. Frankfurt/Main: Suhrkamp.
- Dempster, A.P., Laird, N.M., Rubin, D.B. (1977). Maximum likelihood estimation from incomplete data via the EM-algorithm. *Journal of the Royal Statistical Society*, 39, 1-22.
- Dörner, D. (1988). *Die Logik des Mißlingens. Strategisches Denken in komplexen Situationen*. Reinbek: Rowohlt.
- Dollard, J., Doob, L.W., Miller, N.E., Mowrer, O.H., Sears, R.R. (1939). *Frustration and Aggression*. New Haven: Yale University Press.
- Dransfeld, K., Kienle, P., Vonach, H. (1992). *Physik I. Newtonsche und relativistische Mechanik. Einführungskurs für Studierende der Naturwissenschaften und Elektrotechnik*. 6. Auflage. München: Oldenbourg.
- Doob, J.L. (1934). Probability and statistics. *Trans. Amer. Math. Soc.*, 36, 759-775.
- Dugué, D. (1937). Application des propriétés de la limite au sens du calcul des probabilités à l'étude des diverses questions d'estimation. *J. École Polytechnique*, 1937, 305-373.
- Eysenck, H.J. (1985). *Sigmund Freud: Niedergang und Ende der Psychoanalyse*. München: List.
- Fischer, G.H. (Hg.) (1968). *Psychologische Testtheorie*. Bern: Huber.
- Fischer, G.H. (1974). *Einführung in die Theorie psychologischer Tests*. Bern: Huber.
- Fischer, G.H. (1974b). Lineare logistische Modelle zur Beschreibung von Einstellungs- und Verhaltensänderungen unter dem Einfluss von Massenkommunikation. In: Kempf, W. (Hg.). *Probabilistische Modelle in der Sozialpsychologie*. Bern: Huber, 81-127.
- Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Phil. Trans. Royal Soc.*, 222, 309-368.

- Fisher, R.A. (1925). Theory of statistical estimation. Proceedings of the Cambridge Philosophical Society, 22, 700-725
- Fisher, R.A. (1935). The design of experiments. Edinburgh: Oliver & Boyd.
- Fisz, M. (1970). Wahrscheinlichkeitsrechnung und mathematische Statistik. Berlin: VEB Deutscher Verlag der Wissenschaften.
- Flick, U., Kardorff, E.v., Steinke, I. (Hg.) (2000). Qualitative Sozialforschung. Ein Handbuch. Reinbek: Rowohlt.
- Foerster, H.v. (1972). Responsibility of competence. Journal of Cybernetics, 2/2. In: Foerster, H.v. (2001). Short Cuts. Frankfurt/Main: Zweitausendeins.
- Foerster, H.v. (1994). Wissenschaft des Unwißbaren. In: Fedrowitz, D., Matejowski, D., Kaiser, G. (Hg.). Neuroworlds. Gehirn-Geist-Kultur. Frankfurt/Main: Campus.
- Foppa, K. (1965). Lernen, Gedächtnis, Verhalten. Ergebnisse und Probleme der Lernpsychologie. Köln: Kiepenheuer & Witsch.
- Franck, D. (1985). Verhaltensbiologie. 2. Auflage. Stuttgart: dtv-Thieme.
- Frege, G. (1879). Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle: Nebert.
- Freud, S. (1900). Die Traumdeutung. Studienausgabe, Band 2. Frankfurt/Main: Fischer (1972).
- Freud, S. (1917). Vorlesungen zur Einführung in die Psychoanalyse. Gesammelte Werke, Band 11. Frankfurt/Main: Fischer (1980).
- Freud, S. (1937). Konstruktionen in der Analyse. Studienausgabe, Ergänzungsband: Schriften zur Behandlungstechnik. Frankfurt/Main: Fischer (1975), 393-406.
- Fromm, E. (1979). Sigmund Freuds Psychoanalyse – Größe und Grenzen. Gesamtausgabe, Band 8. Stuttgart: Deutsche Verlags-Anstalt (1981), 259-362.
- Fuchs, A. (1993). Gewaltbegriff und Funktion von Gewalt. In: Kempf, W., Frindte, W., Sommer, G., Spreiter, M. (Hg.). Gewaltfreie Konfliktlösungen. Heidelberg: Asanger, 35-52.
- Fuchs, A., Maihöfer, J. (1992). Gewalttätig sind vor allem die anderen. Perspektivendivergenz der Gewaltzuschreibung als Problem einer empirischen Bestimmung des Gewaltbegriffs. In: Präsidentin der Humboldt-Universität zu Berlin (Hg.). Friedenspsychologie im Spannungsfeld zwischen Psychologie und Politischer Wissenschaft. Berlin: Humboldt-Universität.
- Galtung, J. (1969). Gewalt, Frieden und Friedensforschung. In: Galtung, J. Strukturelle Gewalt. Reinbek: Rowohlt Taschenbuchverlag (1975).
- Galtung, J. (1993). Friedensforschung als universitäres Studienfach: wie geht es weiter? In: Alfs, M., Dominikowski, T. Hiegemann, M., Kinkelbur, D., Nabers, D., Westphal, N. (Hg.). Arbeit am verlorenen Frieden. Münster: Agenda.

- Gauß, C.F. (1880). Werke. Band 4. Göttingen: Königl. Ges. der Wissenschaften.
- Geertz, C. (1987). Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme. Frankfurt/Main: Suhrkamp.
- Gigerenzer, G., Murray, D.J. (1987). Cognition as intuitive statistics. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Gill, A. (1962). Introduction to the theory of finite state machines. New York: McGraw-Hill.
- Glaser, B.G., Strauss, A.L. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine Publ. Co.
- Glaser, W.R. (2000). Systemtheorie. In: Straub, J., Kempf, W., Werbik, H. (Hg.). Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven. München: Deutscher Taschenbuch Verlag, 68-96.
- Glasl, F. (1992). Konfliktmanagement. Ein Handbuch zur Diagnose und Behandlung von Konflikten für Organisationen und ihre Berater. Bern: Haupt.
- Gödel, K. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalküls. Mh. Math. Phys., 37, 349-360.
- Goodman, L.A. (1974). The analysis of systems of qualitative variables when some of the variables are unobservable, Part I. A modified latent structure approach. American Journal of Sociology, 79, 1179-1259.
- Goodman, L.A. (1979). On the estimation of parameters in latent structure analysis. Psychometrika, 44, 123-128.
- Gormezano, I., Moore, J.W. (1969). Classical conditioning. In: Marx, M.H. (ed.). Learning: Processes. New York: Macmillan, 119-203.
- Groeben, N., Erb, E. (2000). Menschenbilder. In: Straub, J., Kempf, W., Werbik, H. (Hg.). Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven. München: Deutscher Taschenbuch Verlag, 17-41.
- Groeben, N., Nüse, R., Gauler, E. (1992). Diagnose argumentativer Unintegrität. Zeitschrift für experimentelle und angewandte Psychologie, 39, 533-558.
- Grubitzsch, S., Rexilius, G. (Hg.) (1978). Testtheorie und Testpraxis. Reinbek: Rowohlt.
- Grünbaum, A. (1988). Die Grundlagen der Psychoanalyse: Eine philosophische Kritik. Stuttgart: Reclam.
- Gulliksen, H. (1950). Theory of mental tests. New York: Wiley.
- Guttman, L. (1950). The basis of scalogram analysis. In: Stouffer, S.A., Guttman, L., Suchman, E.A., Lazarsfeld, P.F., Star, S.A., Clausen, J.A. (eds.). Studies in social psychology in world war II, Vol. IV. Measurement and prediction. Princeton, N.J.: Princeton University Press, 60-90.
- Habermas, J. (1968). Erkenntnis und Interesse. Frankfurt/Main: Suhrkamp.

- Hartmann, D., Werbik, H. (2001). Über Reichweite und Grenzen einer naturwissenschaftlichen Psychologie. *Handlung, Kultur, Interpretation*, 10/1, 158-179.
- Heinze, T., Klusemann, H.W. (1980). Versuch einer sozialwissenschaftlichen Paraphrasierung am Beispiel des Ausschnittes einer Bildungsgeschichte. In: Heinze, Th., Klusemann, H.W., Soeffner, H.G. (Hg.). *Interpretationen einer Bildungsgeschichte. Überlegungen zur sozialwissenschaftlichen Hermeneutik*. Bensheim: päd.-extra-Buchverlag, 97-152.
- Hempel, C.G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.
- Hempel, C.G., Oppenheim, P. (1948). *Studies in the logic of explanation*. *Philosophy of Science*, 15, 135-175.
- Hoffmann-Riem, Ch. (1980). Die Sozialforschung einer interpretativen Soziologie. *Der Datengewinn. Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 32, 339-372.
- Holt, E.B. (1931). *Animal drive and the learning process. An essay toward radical empiricism*. London: Williams & Norgate.
- Holzkamp, K. (1965). Zur Problematik der Realitätsverdoppelung in der Psychologie. *Psychologische Rundschau*, 16, 209-222.
- Holzkamp, K. (1972). *Kritische Psychologie*. Frankfurt/Main: Fischer.
- Holzkamp, K. (1986). Die Verkennung von Handlungsbegründungen als empirische Zusammenhangsbehauptungen in sozialpsychologischen Theorien: Methodologische Fehlorientierung infolge von Begriffsverwirrung. *Zeitschrift für Sozialpsychologie*, 17, 216-238.
- Holzkamp, K. (1994). Am Problem vorbei. Zusammenhangsblindheit der Variablenpsychologie. *Forum Kritische Psychologie* 34, 80-94.
- Hopf, C. (2000). Qualitative Interviews – ein Überblick. In: Flick, U., Kardorff, E.v., Steinke, I. (Hg.). *Qualitative Sozialforschung. Ein Handbuch*. Reinbek: Rowohlt, 349-359.
- Hotelling, H. (1930). The consistency and ultimate distribution of optimum statistics. *Trans. Amer. Math. Soc.*, 32, 847-859.
- Hoyningen-Huene, P. (1983). Autonome historische Prozesse – kybernetisch betrachtet. *Geschichte und Gesellschaft*, 9/1, 119-123.
- Hoyningen-Huene, P. (1989). *Die Wissenschaftsphilosophie Thomas S. Kuhns. Rekonstruktion und Grundlagenprobleme*. Braunschweig: Vieweg.
- Hoyningen-Huene, P. (1989b). Naturbegriff – Wissensideal – Experiment. Warum ist die neuzeitliche Naturwissenschaft technisch verwertbar? *Zeitschrift für Wissenschaftsforschung*, 5, 43-55.
- Hoyningen-Huene, P. (1997). Der Empiriebezug der Psychologie: Philosophische Bemerkungen zur Position Jochen Brandtstädters. *Zeitschrift für Sozialpsychologie*, 28, 229-240.
- Huber, O. (1987). *Das psychologische Experiment: Eine Einführung*. Bern: Huber.

- Hull, C.L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.
- Inhelder, B., Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Janich, P. (1980). Behaviorismus. In: Mittelstraß, J. (Hg.). *Enzyklopädie Philosophie und Wissenschaftstheorie*. Band 1. Mannheim: Bibliographisches Institut.
- Janich, P. (1995). Prototheorie. In: Mittelstraß, J. (Hg.). *Enzyklopädie Philosophie und Wissenschaftstheorie*. Band 3. Stuttgart: Metzler.
- Janich, P., Kambartel, F., Mittelstraß, J. (1974). *Wissenschaftstheorie als Wissenschaftskritik*. Frankfurt/Main: Aspekte Verlag.
- Jones, J.E. (1962). Contiguity and reinforcement in relation to CS-UCS intervals in classical aversive conditioning. *Psychological Review*, 69, 176-186.
- Kahnemann, D., Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.
- Kambartel, F. (1968a). Was ist und soll Philosophie? In: Kambartel, F. *Theorie und Begründung*. Frankfurt/Main: Suhrkamp (1976), 11-27.
- Kambartel, F. (1968b). Was ist Logik? In: Kambartel, F. *Theorie und Begründung*. Frankfurt/Main: Suhrkamp (1976), 95-110.
- Kambartel, F. (1981). Friede und Konfliktbewältigung, handlungstheoretisch und ethisch betrachtet. In: Kempf, W., Aschenbach, G. (Hg.). *Konflikt und Konfliktbewältigung*. Bern: Huber, 181-194.
- Kamlah, W., Lorenzen, P. (1967). *Logische Propädeutik*. Mannheim: Bibliographisches Institut.
- Kelle, U., Erzberger, Ch. (2000). Qualitative und quantitative Methoden: kein Gegensatz. In: Flick, U., Kardoff, E.V., Steinke, I. (Hg.). *Qualitative Forschung. Ein Handbuch*. Reinbek: Rowohlt, 299-308.
- Kempf, W. (1972). Probabilistische Modelle experimentalpsychologischer Versuchssituationen. *Psychologische Beiträge*, 14, 16-37.
- Kempf, W. (1978). *Konfliktlösung und Aggression. Zu den Grundlagen einer psychologischen Friedensforschung*. Bern: Huber.
- Kempf, W. (1986). Zur Aufgabenorientiertheit von Wissenschaft und der Funktion von Theorie im traditionellen und interpretativen Paradigma. In: Amelang, M. (Hg.). *Bericht über den 35. Kongreß der Deutschen Gesellschaft für Psychologie in Heidelberg 1986*. Band 1. Göttingen: Hogrefe, 51-52.
- Kempf, W. (1987). Psychologische Forschung als Begegnung. In: Brandtstädter, J. (Hg.). *Struktur und Erfahrung in der psychologischen Forschung*. Berlin: de Gruyter, 169-207.
- Kempf, W. (1993). Konflikteskalation durch autonome Prozesse. In: Kempf, W., Frindte, W., Sommer, G., Spreiter, M. (Hg.). *Gewaltfreie Konfliktlösungen*. Heidelberg: Asanger, 53-70.
- Kempf, W. (1996). Begriff und Probleme des Friedens. *Beiträge der Sozialpsychologie. Kurseinheit 1: Aggression, Gewalt und Gewaltfreiheit*. Hagen: Fernuniversität.

- Kempf, W. (1999). Zum Verhältnis von Allgemeinem und Konkretem in der subjektwissenschaftlichen Forschung. In: Straub, J., Werbik, H. (Hg.). *Handlungstheorie. Begriff und Erklärung des Handelns im interdisziplinären Diskurs*. Frankfurt/Main: Campus, 93-110.
- Kempf, W. (2000). Gewaltursachen und Gewaltdynamiken. In: ÖSFK (Hg.). *Konflikt und Gewalt. Ursachen – Entwicklungstendenzen – Perspektiven*. Münster: Agenda, 44-65.
- Kempf, W. (2003). Theoretical and empirical foundations. In: Austrian Study Center for Peace and Conflict Resolution (ed.). *Constructive conflict coverage. A social psychological approach*. Berlin: regener.
- Klahr, D. (1978). Information processing models of cognitive development. In: Scandura, J.M., Brainerd, Ch.J. (eds.). *Structural / process models of complex human behavior*. Alphen aan den Rijn: Sijthoff & Noordhoff, 479-518.
- Klahr, D. & Siegler, R.S. (1978). The representation of children's knowledge. In: Siegler, R.S. (ed.). *Children's thinking: what develops?* New York: Erlbaum.
- Klatzky, R.L. (1980). *Human memory, structures and processes*. San Francisco: Freeman.
- Kochinka, A., Werbik, H. (2000). Logische Propädeutik und Wissenschaftstheorie. In: Straub, J., Kempf, W., Werbik, H. (Hg.). *Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven*. München: Deutscher Taschenbuch Verlag, 42-67.
- Kolb, J. (1993). Rekonstruktion theoretischer Datenstrukturen mittels Latent-Class-Analyse – zur Eignung des Modells als theoriengenerierendes Verfahren. Universität Konstanz: Diplomarbeit.
- Kolmogoroff, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Springer.
- Kracauer, S. (1952). The challenge of qualitative content analysis. *Public Opinions Quarterly*, 16, 631-642.
- Kraus, W. (2001). Text. In: Keupp, H., Weber, K. (2001). *Psychologie. Ein Grundkurs*. Reinbek: Rowohlt.
- Kreyszig, E. (1968). *Statistische Methoden und ihre Anwendungen*. 3. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Lamnek, S. (1988). *Qualitative Sozialforschung. Band 1. Methodologie*. Weinheim: Psychologie Verlags Union.
- Lamnek, S. (1989). *Qualitative Sozialforschung. Band 2. Methoden und Techniken*. Weinheim: Psychologie Verlags Union.
- Laplanche, J., Pontalis, J.-B. (1973). *Das Vokabular der Psychoanalyse*. Frankfurt/Main: Suhrkamp.
- Lazarsfeld, P.F. (1950). Logical and mathematical foundations of latent structure analysis. In: Stouffer, S.A., Guttman, L., Suchman, E.A., Lazarsfeld, P.F., Star, S.A., Clausen, J.A. (eds.). *Studies in social psychology in world war II, Vol. IV. Measurement and prediction*. Princeton/N.J.: Princeton University Press, 362-412.

- Lazarsfeld, P.F. (1959). Latent structure analysis. In: Koch, S. (ed.). *Psychology: a study of a science*. Vol. 3. New York: McGraw-Hill, 476-542.
- Lazarsfeld, P.F. (1961). The algebra of dichotomous systems. In: Solomon, H. (ed.). *Studies in item analysis and prediction*. Stanford: Stanford University Press, 111-157.
- Lazarsfeld, P.F., Henry, N.W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin Co.
- Leithäuser, T., Volmerg, B. (1977). Die Entwicklung einer empirischen Forschungsperspektive aus der Theorie des Alltagsbewusstseins. Leithäuser, T., Volmerg, B., Salje, G., Volmerg, U., Wutka, B. *Entwurf zu einer Empirie des Alltagsbewußtseins*. Frankfurt/Main: Suhrkamp, 11-159.
- Lewin, K. (1931). Der Übergang von der aristotelischen zur galileischen Denkweise in Biologie und Psychologie. *Erkenntnis*, 1, 421-466.
- Lord, F.M. (1952). A theory of test scores. *Psychometric Monograph*, No. 7.
- Lord, F.M. (1953a). An application of confidence intervals and of maximum likelihood to the estimation of an examinee's ability. *Psychometrika*, 18, 57-75.
- Lord, F.M. (1953b). The relation of test score to the trait underlying the test. *Educational and Psychological Measurement*, 13, 517-548.
- Lord, F.M., Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley.
- Lorenz, Konrad (1963, 1974). *Das sogenannte Böse. Zur Naturgeschichte der Aggression*. Wien: Borotha-Schoeler. München: Deutscher Taschenbuchverlag.
- Lorenz, Kuno (1984). Klassische Logik. In: Mittelstraß, J. (Hg.). *Enzyklopädie Philosophie und Wissenschaftstheorie*. Band 2. Mannheim: Bibliographisches Institut.
- Lorenz, Kuno (1996). Wahrheit. In: Mittelstraß, J. (Hg.). *Enzyklopädie Philosophie und Wissenschaftstheorie*. Band 4. Stuttgart: Metzler.
- Lorenzen, P. (1974). Zur Konstruktion von "Wahrscheinlichkeit". In: Lorenzen, P. *Konstruktive Wissenschaftstheorie*. Frankfurt/Main: Suhrkamp, 209-218.
- Lorenzen, P. (1985). Eine konstruktive Deutung des Dualismus in der Wahrscheinlichkeitstheorie. In: Lorenzen, P. *Grundbegriffe technischer und politischer Kultur*. Frankfurt/Main: Suhrkamp, 59-84.
- Lorenzen, P. (1987). *Lehrbuch der konstruktiven Wissenschaftstheorie*. Mannheim: Bibliographisches Institut.
- Lorenzen, P., Schwemmer, O. (1975). *Konstruktive Logik, Ethik und Wissenschaftstheorie*. Mannheim: Bibliographisches Institut.
- Lüders, L. (2000). Beobachten im Feld und Ethnographie. In: Flick, U., Kardorff, E.v., Steinke, I. (Hg.). *Qualitative Sozialforschung. Ein Handbuch*. Reinbek: Rowohlt, 384-401.

- Mackintosh, N.J. (1974). *The psychology of animal learning*. New York: Academic Press.
- Markl, H. (1982). Evolutionsbiologie des Aggressionsverhaltens. In: Hilke, R., Kempf, W. (Hg.). *Aggression. Naturwissenschaftliche und kulturwissenschaftliche Perspektiven der Aggressionsforschung*. Bern: Huber, 21-43.
- May, R. (1979). *Wie entwickelt sich das Verständnis von Proportionalität? Strukturelle Modelle zum Problemlösen bei Balkenwaage-Aufgaben*. Universität Konstanz: Diplomarbeit.
- McDougall, W. (1912). *Psychology. The study of behavior*. London: Williams and Norgate.
- Meehl, P.E. (1950). On the circularity of the law of effect. *Psychological Bulletin*, 47, 52-75.
- Meier, C. (1978). Fragen und Thesen zu einer Theorie historischer Prozesse. In: Faber, K.G., Meier, C. (Hg.). *Historische Prozesse*. München: Deutscher Taschenbuchverlag, 11-66.
- Meinefeld, W. (2000). Hypothesen und Vorwissen in der qualitativen Sozialforschung. In: Flick, U., Kardorff, E.v., Steinke, I. (Hg.). *Qualitative Sozialforschung. Ein Handbuch*. Reinbek: Rowohlt, 265-275.
- Mertens, W. (1991). *Einführung in die psychoanalytische Therapie*. Band 3. Stuttgart: Kohlhammer.
- Meschkowski, H. (1966). *Grundlagen der euklidischen Geometrie*. Mannheim: Bibliographisches Institut.
- Meschkowski, H. (1986). *Mathematik verständlich dargestellt*. 3. Auflage. München: Piper.
- Micko, H.C. (1970). Eine Verallgemeinerung der Messmodells von Rasch mit einer Anwendung auf die Psychophysik der Reaktionen. *Psychologische Beiträge*, 12, 4-22.
- Miller, N.E. (1941). The frustration-aggression-hypothesis. In: Berkowitz, L. (ed.). *Roots of aggression*. New York: Atherton Press, 29-34.
- Milgram, S. (1974). *Das Milgram-Experiment*. Reinbek: Rowohlt.
- Mittelstraß, J. (1996). Szientismus. In: Mittelstraß, J. (Hg.). *Enzyklopädie Philosophie und Wissenschaftstheorie*. Band 4. Stuttgart: Metzler, 194-195.
- Mollenhauer, K., Rittelmeyer, C. (1977). *Methoden der Erziehungswissenschaft*. München: Juventa.
- Morgan, C.T. (1961). *Introduction to psychology*. New York: McGraw-Hill.
- Müller, B., Büttner, C. (1996). Optimierungschancen von Peacekeeping, Peacemaking und Peacebuilding durch gewaltfreie Interventionen? Arbeitspapier 4. Wahlenau: Institut für Friedensarbeit und gewaltfreie Konfliktaustragung.
- Murdock, G.P. (1962). *Ethnographic Atlas*. *Ethnology*, 1, 113-134.
- Murdock, G.P. (1963). *Ethnographic Atlas*. *Ethnology*, 2, 159-253.
- Murdock, G.P. (1964). *Ethnographic Atlas*. *Ethnology*, 3, 199-217.

- Neidhard, F. (1986). Gewalt. Soziale Bedeutungen und sozialwissenschaftliche Bestimmung des Begriffs. In: Bundeskriminalamt (Hg.). Was ist Gewalt? Wiesbaden: Bundeskriminalamt.
- Neyman, J. (1935). Baltimore, Md.: Waverly Pr, Journal of the Institute of Mathematical Statistics. *Annals of Mathematical Statistics*, 6, 111-116.
- Neyman, J. (1938). L'estimation statistique traitée comme un problème classique de probabilité. *Actualités Scientifiques et Industrielles*, 739, 25-57.
- Neyman, J., Pearson, E.S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. Part I. *Biometrika*, 20A, 175-240.
- Neyman, J., Pearson, E.S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. Part II. *Biometrika*, 20A, 263-294.
- Neyman, J., Scott, E.L. (1948). Consistent estimates based on partially consistent observations. *Econometrica*, 16, 1-32.
- Novick, M. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3, 1-18.
- Orth, B. (1974). Einführung in die Theorie des Messens. Stuttgart: Kohlhammer.
- Otterbein, K.F. (1968). Internal war: A cross-cultural study. *American Anthropologist*, 70, 277-289.
- Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50, 157-175.
- Pfanzagl, J. (1968). Theory of measurement. Würzburg: Physica.
- Pfanzagl, J. (1993). A case of asymptotic equivalence between conditional and marginal maximum likelihood estimators. *Journal of Statistical Planning and Inference*, 35, 301-307.
- Popper, K. (1994). Logik der Forschung. 10. Auflage. Tübingen: Mohr.
- Postman, L. (1947). The history and present status of the law of effect. *Psychological Bulletin*, 44, 489-563.
- Rao, C.R. (1945). Information and accuracy obtainable in an estimation of a statistical parameter. *Bulletin of the Calcutta Mathematical Society*, 37, 81-91.
- Rasa, O.A.E. (1969). The effect of pair isolation on reproductive success in *Etroplus maculatus* (Cichlidae). *Zeitschrift für Tierpsychologie*, 26, 846-882.
- Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen: Nielsen & Lydiche.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. Berkeley: University of California Press.

- Rasch, G. (1968). A mathematical theory of objectivity and its consequences for model construction. Proceedings of the European Meeting on Statistic, Econometrics and Management Science, Amsterdam, September 2-7, 1968.
- Rohracher, H. (1963). Einführung in die Psychologie. Achte Auflage. Wien: Urban & Schwarzenberg.
- Reimann, M. (2002). Communication disorders in conflict coverage. In: Kempf, W., Luostarinen, H. (eds.). Journalism and the New World Order. Vol. 2. Studying war and the media. Göteborg: Nordicom.
- Rogers, C.R. (1972). Die klientenzentrierte Gesprächspsychotherapie. München: Kinkler.
- Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. Applied Psychological Measurement, 14, 271-282.
- Rost, J. (1996). Lehrbuch Testtheorie Testkonstruktion. Bern: Huber.
- Rost, J. (1997). Intelligenz. In: Straub, J., Kempf, W., Werbik, H. (Hg.). Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven. München: Deutscher Taschenbuchverlag, 530-558.
- Schneirla, T.C. (1959). An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In: Jones, M.R. (ed.). Nebraska Symposium on Motivation, Vol. 7. Lincoln: University of Nebraska Press, 1-42.
- Schütze, F. (1977). Die Technik des narrativen Interviews in Interaktionsfeldstudien – dargestellt an einem Projekt zur Erforschung von kommunalen Machtstrukturen. Universität Bielefeld, Fakultät für Soziologie: Arbeitsberichte und Forschungsmaterialien, Nr. 1.
- Schütze, F. (1978). Was ist "kommunikative Sozialforschung"? In: Gärtner, A., Hering, S. (eds.). Modellversuch "Soziale Studiengänge an der GH Kassel". Materialien 12: Regionale Sozialforschung. Kassel: Gesamthochschule.
- Schulz von Thun, F. (1981). Miteinander reden 1. Störungen und Klärungen. Allgemeine Psychologie der Kommunikation. Reinbek: Rowohlt.
- Schwemmer, O. (1976). Theorie der rationalen Erklärung. München: Beck.
- Schwemmer, O. (1987). Das Allgemeine unseres Handelns: Zum Sinn der Rede von Gesetzen und Regeln für die Darstellung menschlichen Handelns. In: Brandtstädter, J. (Hg.). Struktur und Erfahrung in der psychologischen Forschung. Berlin: de Gruyter, 104-124.
- Sears, R.R. (1941). Nonaggressive reactions to frustration. Psychological Review, 48, 343-346.
- Selg, H. (1975). Einführung in die experimentelle Psychologie. 4. Auflage. Stuttgart: Kohlhammer.
- Sherif, M., Sherif, C.W. (1969). Social Psychology. New York: Harper & Row.

- Siegler, R.S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, 8, 481-520.
- Sipes, R.G., 1973. War, sports and aggression. *American Anthropologist*, 75, 64-86.
- Skinner, B.F. (1953). *Science and human behavior*. New York: Macmillan.
- Smedslund, J. (1976). A reexamination of the role of theory in psychology. Vortrag am 21. Internationalen Kongreß für Psychologie, Paris, 18.-25. Juli 1976; hektographiert.
- Smedslund, J. (1988). *Psycho-logic*. New York: Springer.
- Smedslund, J. (1991). The pseudoempirical in psychology and the case for psycho-logic. *Psychological Inquiry*, 2, 325-338.
- Sokolov, E.N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, 25, 545-580.
- Spada, H., Ernst, A.M., Ketterer, W. (1990). Klassische und operante Konditionierung. In: Spada, H. (Hg.). *Lehrbuch Allgemeine Psychologie*. Bern: Huber.
- Spearman, Ch. (1904). "General intelligence", objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Spetch, M.L., Wilkie, D.M., Pinel, L.P.J. (1981). Backward conditioning: A reevaluation of the empirical evidence. *Psychological Bulletin*, 89, 163-175.
- Stern, W. (1920). *Die Intelligenz der Kinder und Jugendlichen und die Methoden ihrer Untersuchung*. Leipzig: Barth.
- Steyer, R., Eid, M. (1993). *Messen und Testen*. Ein Lehrbuch. Berlin: Springer.
- Straßmaier, S. (2003). *Analyse der Psychoanalyse*. Der Kritische Rationalismus und die Widerstandskonzepte Sigmund Freuds. Berlin: regner.
- Straub, J. (1999). Narrative Handlungserklärungen im Lichte einer pragmatisch-epistemischen Erklärungstheorie. In: Straub, J., Werbik, H. (Hg.). *Handlungstheorie. Begriff und Erklärung des Handelns im interdisziplinären Diskurs*. Frankfurt/Main: Campus, 261-284.
- Suppes, P., Zinnes, J.L. (1963). Basic measurement theory. In: Luce, R., Bush, R.R., Galanter, E. (eds.). *Handbook of mathematical psychology*. New York: Wiley, 1-76.
- Therstappen, A. (1980). *Immunisierungsstrukturen bei Sigmund Freud: Ein Beitrag zur wissenschaftstheoretischen Auseinandersetzung mit Freuds Entwurf der Psychoanalyse*. Unveröffentlichte Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen.
- Thissen, D. (1982). Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika*, 47, 175-186.
- Thomä, H., Kächele, H. (1985). *Lehrbuch der psychoanalytischen Therapie*. 1 Grundlagen. Berlin: Springer.
- Thorndike, E.L. (1913). *The psychology of learning*. New York: Teachers College.

- Thurstone, L.L., Thurstone, T.G. (1941). *Factorial studies of intelligence*. Chicago: The University of Chicago Press.
- Tolman, E.C. (1937). The acquisition of string-pulling by rats – conditioned response or sign-gestalt? *Psychological Review*, 44, 195-211.
- Turing, A. M. (1936/37). On Computable Numbers with an Application to the Entscheidungsproblem. *Proc. Lond. Math. Soc. (ser. 2)*, 42, 230-265 und 43, 544-546.
- Tschebyscheff, P.L. (1948). *О средних величинах. Полное собрание сочинений*, т. 2. Москва – Ленинград.
- Volmerg, B., Volmerg, U., Leithäuser, T. (1983). *Kriegsängste und Sicherheitsbedürfnis. Zur Sozialpsychologie des Ost-West-Konflikts im Alltag*. Frankfurt/Main: Fischer Taschenbuchverlag.
- Volmerg, U. (1977). Kritik und Perspektiven des Gruppendiskussionsverfahrens in der Forschungspraxis. In: Leithäuser, T., Volmerg, B., Salje, G., Volmerg, U., Wutka, B. *Entwurf zu einer Empirie des Alltagsbewußtseins*. Frankfurt/Main: Suhrkamp, 184-217.
- Wald, A. (1943). An extension of Wilk's method for setting tolerance limits. *Annals of Math. Statistics*, 14, 45-55.
- Waldenfels, B. (1999). Symbolik, Kreativität und Responsivität. Grundzüge einer Phänomenologie des Handelns. In: Straub, J., Werbik, H. (Hg.). *Handlungstheorie. Begriff und Erklärung des Handelns im interdisziplinären Diskurs*. Frankfurt/Main: Campus, 243-260.
- Warm, T.A. (1989). Weighted likelihood estimation of ability in item response models. *Psychometrika*, 54, 427-450.
- Watson, J.G. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20, 158-177.
- Watson, J.G. (1930). *Behaviorism*. New York: Norton.
- Watzlawick, P., Beavin, J.K., Jackson, D.D. (1967). *Pragmatics of human communication. A study of interactional patterns, pathologies, and paradoxes*. New York: Norton & Company.
- Watzlawick, P., Beavin, J.K., Jackson, D.D. (1969). *Menschliche Kommunikation. Formen, Störungen, Paradoxien*. Bern: Huber.
- Werbik, H. (1971). Das Problem der Definition "aggressiver" Verhaltensweisen. *Zeitschrift für Sozialpsychologie*, 2, 233-247.
- Werbik, H. (1974). *Theorie der Gewalt*. München: UTB W. Fink.
- Werbik, H. (1982). Zur terminologischen Bestimmung von Aggression und Gewalt. In: Hilke, R., Kempf, W. (Hg.). *Aggression. Naturwissenschaftliche und kulturwissenschaftliche Perspektiven der Aggressionsforschung*. Bern: Huber, 334-350.
- Werbik, H., Appelsmeyer, H. (1999). Handlungsbegriff und Perspektivität. Eine Diskussion traditioneller Perspektiven psychologischer Erkenntnisgewinnung. In: Straub, J., Werbik, H. (Hg.). *Handlungstheorie. Begriff und Erklärung des Handelns im interdisziplinären Duskurs*. Frankfurt/Main: Campus, 74-92.

- Westmeyer, H. (1973). Kritik der psychologischen Unvernunft. Probleme der Psychologie als Wissenschaft. Stuttgart: Kohlhammer.
- Wiener, N. (1968). Kybernetik. Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine. Reinbek: Rowohlt.
- Wittgenstein, L. (1953). Philosophische Untersuchungen. Oxford: Blackwell.
- Witzel, A. (1982). Verfahren der qualitativen Sozialforschung. Überblick und Alternativen. Frankfurt/Main: Campus.
- Whorf, B.L. (1956). Science and linguistics. In: Carroll, J.B. (ed.). Language, thought and reality. Writings of Benjamin Lee Whorf. New York: Wiley, 207-219.
- Wolfowitz, J. (1949). On Wald's proof of the consistency of the maximum likelihood estimate. *Annals of Math. Statistics*, 20, 601-602.
- Wright, G.H.v. (1974). Erklären und Verstehen. Frankfurt/Main: Fischer Athenäum.

Anhang

Über den Autor

Wilhelm Kempf, Prof. Dr. phil. habil., wurde 1947 in Klagenfurt/Österreich geboren. 1965-1970 Studium der Psychologie, Philosophie und Statistik an der Universität Wien. 1969/70 Postgraduiertenstudium der Soziologie am Institut für Höhere Studien. 1970 Promotion an der Universität Wien. 1977 Habilitation an der Universität Erlangen-Nürnberg, *venia legendi* für Psychologie.

Seit 1977 Professor für psychologische Methodenlehre und Leiter der Projektgruppe Friedensforschung an der Universität Konstanz. Lehr- und Forschungsaufenthalte an der Universidad de Costa Rica, an der Universität Tampere (Finnland) und an der York University (Toronto, Kanada).

Arbeitsschwerpunkte: Psychologische Methodologie, Friedens- und Konfliktforschung.

Aktuelle Buchpublikationen: *Los Medios y la Cultura de Paz* (mit Sonia Gutiérrez, 2001); *Journalism and the New World Order, Vol. II. Studying War and the Media* (mit Heikki Luostarinen, 2002); *Constructive Conflict Coverage – A Social Psychological Approach* (herausgegeben vom Austrian Study Center for Peace and Conflict Resolution, 2003).

Adresse: Universität Konstanz, FB Psychologie
D-78457 Konstanz

eMail: Wilhelm.Kempf@uni-konstanz.de

Website: www.uni-konstanz.de/FuF/SozWiss/fg-psy/ag-meth/

Statistische Tabellen

1 Standardnormalverteilung

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{u^2}{2}} du$$

Werte von z für									
prob($Z \leq z$) = $\Phi(z)$ =					prob($Z \geq z$) = $1 - \Phi(z)$ =				
0.001	0.005	0.01	0.025	0.05	0.05	0.025	0.01	0.005	0.001
-3.090	-2.576	-2.326	-1.960	-1.645	1.645	1.960	2.326	2.576	3.090

Werte von z für		
prob($-z \leq Z \leq z$) = $\Phi(z) - \Phi(-z)$ =		
0.95	0.99	0.999
1.960	2.576	3.291

2 χ^2 -Verteilung

$$F(x) = K_n \int_0^x u^{(n-2)/2} e^{-u/2} du \quad \text{für } x \geq 0$$

Werte von x für						Werte von x für					
df	prob($X \geq x$) = 1 - F(x) =					df	prob($X \geq x$) = 1 - F(x) =				
	0.05	0.025	0.01	0.005	0.001		0.05	0.025	0.01	0.005	0.001
1	3.84	5.02	6.63	7.88	10.83	23	35.17	38.08	41.64	44.18	49.73
2	5.99	7.38	9.21	10.60	13.82	24	36.42	39.36	42.98	45.56	51.18
3	7.81	9.35	11.34	12.84	16.27	25	37.65	40.65	44.31	46.93	52.62
4	9.49	11.14	13.28	14.86	18.47	26	38.89	41.92	45.64	48.29	54.05
5	11.07	12.83	15.09	16.75	20.52	27	40.11	43.19	46.96	49.64	55.48
6	12.59	14.45	16.81	18.55	22.46	28	41.34	44.46	48.28	50.99	56.89
7	14.07	16.01	18.48	20.28	24.32	29	42.56	45.72	49.59	52.34	58.30
8	15.51	17.53	20.09	21.96	26.13	30	43.77	46.98	50.89	53.67	59.70
9	16.92	19.02	21.67	23.59	27.88	40	55.76	59.34	63.69	66.77	73.40
10	18.31	20.48	23.21	25.19	29.59	50	67.50	71.42	76.15	79.49	86.66
11	19.68	21.92	24.73	26.76	31.26	60	79.08	83.30	88.38	91.95	99.61
12	21.03	23.34	26.22	28.30	32.91	70	90.53	95.02	100.43	104.22	112.32
13	22.36	24.74	27.69	29.82	34.53	80	101.88	106.63	112.33	116.32	124.84
14	23.68	26.12	29.14	31.32	36.12	90	113.15	118.14	124.12	128.30	137.21
15	25.00	27.49	30.58	32.80	37.70	100	124.34	129.56	135.81	140.17	149.45
16	26.30	28.85	32.00	34.27	39.25	Näherung für df > 100					
17	27.59	30.19	33.41	35.72	40.79	$x = \frac{1}{2}(\sqrt{2df - 1} + z)^2$					
18	28.87	31.53	34.81	37.16	42.31	Werte von z für					
19	30.14	32.85	36.19	38.58	43.82	prob($X \geq x$) = 1 - F(x) =					
20	31.41	34.17	37.57	40.00	45.32	0.05	0.025	0.01	0.005	0.001	
21	32.67	35.48	38.93	41.40	46.80	1.645	1.960	2.326	2.576	3.090	
22	33.92	36.78	40.29	42.80	48.27						

Aufgabenlösungen

Zu Kapitel 2.1:

1. $x_{vt} = 51$

Zu Kapitel 2.4:

1. a) Nein.
b) Vp 8 zeigt das Verhalten D. Ihr wird aber ein anderer Zahlenwert (4) zugeordnet als den Vpn 2 und 6, die das selbe Verhalten zeigen. / Vp 5 zeigt das Verhalten C, ihr wird aber der selbe Zahlenwert (1) zugeordnet wie den Vpn 4 und 9, die ein anderes Verhalten (B) zeigen.
c) 0 / 3 / 2 / 1 / 2 / 3 / 0 / 3 / 1 / 2 / 2 / 2
2. a) Nein.
b) $C \not\sim C \Rightarrow$ Verletzung der Irreflexivität / $(A \not\sim D \wedge D \not\sim A) \Rightarrow$ Verletzung der Assymetrie / $[(B \not\sim C \wedge C \not\sim D) \wedge D \not\sim B] \Rightarrow$ Verletzung der Transitivität.
3. Der Test bildet keine Guttman-Skala. Die Antwortvektoren der Vpn 4 und 8 sind mit den Modellannahmen der Guttman-Skala inkompatibel.
4. b) $-2\ln(\lambda) = 4.0974$, $df = 3$, n.s.
5. a) Gewichte gleich und Distanzen gleich.
b) Distanzen gleich, Gewichte verschieden.
c) Gewichte gleich, Distanzen verschieden.
d) Gewichte und Distanzen gegenläufig. Höheres Gewicht überwiegt größere Distanz.
e) Gewichte und Distanzen gegenläufig. Größere Distanz überwiegt höheres Gewicht..
f) Gewichte und Distanzen gegenläufig. Höheres Gewicht und größere Distanz kompensieren einander.
6. 1) Distanz
2) Konflikt-Balance
3) Konflikt-Distanz
4) Gewicht
5) Konflikt-Gewicht
6) Balance
7. Die zu erwartenden Antworten lauten:

Item Nr.	1	2	3	4	5	6
a) Stufe 1	Balance	Links	Links	Links	Rechts	Balance
b) Stufe 2	Links	Links	Links	Links	Rechts	Balance
c) Stufe 3	Links	?	?	Links	?	Balance
d) Stufe 4	Links	Balance	Rechts	Links	Rechts	Balance

8. Die Lösungswahrscheinlichkeiten betragen:

	Item Nr.	1	2	3	4	5	6
a)	Stufe 1	0	0	0	1	1	1
b)	Stufe 2	1	0	0	1	1	1
c)	Stufe 3	1	0.33	0.33	1	0.33	1
d)	Stufe 4	1	1	1	1	1	1

9. a) 3

b) 4

c) 4

d) 6

10.a) Stufe 1 < Stufe 2 = Stufe 3 < Stufe 4

b) Aufgabe 1 wird erst ab Stufe 2 gelöst. Die Aufgaben 2 und 3 haben ab Stufe 3 eine positive Lösungswahrscheinlichkeit, werden aber erst ab Stufe 4 mit an Sicherheit grenzenden Wahrscheinlichkeit gelöst. Aufgabe 5 wird bereits auf Stufe 1 und 2 richtig beantwortet. Auf Stufe 3 nimmt die Lösungswahrscheinlichkeit dramatisch ab, und erst ab Stufe 4 wird die Aufgabe wieder mit Sicherheit gelöst.

11.a) Nicht.

b) Stufe 2 gibt mit an Sicherheit grenzender Wahrscheinlichkeit eine richtige Antwort auf Aufgabe 5 und eine falsche Antwort auf die Aufgaben 2 und 3.

Auf Stufe 3 haben die Probanden dagegen bei allen diesen Aufgaben die selbe Lösungswahrscheinlichkeit von 0.33.

12.a) $p_1 = 0.25$; $p_2 = 0.75$.

b) Die Wahrscheinlichkeiten der Antwortmuster betragen:

Antwortmuster	(1,0,0)	(1,1,0)	(0,1,0)	(0,0,1)
Klasse 1 (= Stufe 2)	0.729	0.081	0.009	0.009
Klasse 2 (= Stufe 3)	0.147	0.063	0.147	0.147

c) Die Wahrscheinlichkeit der Antwortmatrix beträgt 0.00024988.

13.a) 32

b) 243

c) 72

14. Die Membership-Wahrscheinlichkeiten betragen:

Antwortmuster	(1,0,0)	(1,1,0)	(0,1,0)	(0,0,1)
Klasse 1 (= Stufe 2)	0.623	0.300	0.020	0.020
Klasse 2 (= Stufe 3)	0.377	0.700	0.980	0.980

Diagnostische Entscheidung: Vp 1 wird Klasse 1 (= Stufe 2) zugeordnet, Vpn 2, 3 und 4 werden Klasse 2 (= Stufe 3) zugeordnet.

15.a) 3

b) $-2\ln(\lambda) = 2.058$, $df = 1$, $p > 0.05$ n.s. Die LCA beschreibt die Daten ebenso gut wie das saturierte Modell.

c) $-2\ln(\lambda) = 381.858$, $df = 10$, $p < 0.001$. Daten sind Mischverteilung.

Zu Kapitel 2.6:

1. a) $E(h_n) = 0.7$
b) $\sigma^2(h_n) = 0.0042$
2. Einsetzen von 0.2 für $c\sigma$ und folglich $c = \frac{0.2}{\sigma} = \frac{0.2}{\sqrt{0.0042}} = 3.086$

in die Ungleichung von Tschebyscheff ergibt

$$\text{prob}\{|h_n - p| \geq 0.2\} \leq \frac{1}{c^2} = 0.105$$

3. a) $u_1 = 104.1$; $u_2 = 113.9$
b) $u_1 = 102.56$; $u_2 = 115.44$

Sachregister

- Absolutskala 102, 106, 108, 111, 184
- Abstraktion 228, 330
- Adjunktion 48f., 52, 184, 186, 271
- AIC-Index 131f., 136
- Allgemeine Messtheorie 74, 102-108
- Allquantor 51, 60, 271
- Alternativhypothese 120, 132, 209-216
- Angeborener Auslösemechanismus (AAM) 167
- Antezedens 311f.
- Antwortmatrix 114, 120, 198f.
- Antwortmuster 98, 115, 126-129, 132, 303
- Appellfunktion (von Sprechhandlungen) 272, 274f., 277f., 343f.
- Asch-Versuch 91f.
- Asymmetrie 36, 96, 98, 103
- Aufforderung 32
 - finale 235, 261
 - nonfinale 261
- Aufgabenverständnis (von Wissenschaft) 62ff.
- Ausdrucksfunktion (von Sprechhandlungen) 273-277, 343f.
- Aussage 15, 28, 32f., 38f., 48-57, 78, 106ff., 118, 150, 160-165, 174f., 185, 187, 199, 207, 217, 228, 231ff., 242, 270, 294, 304, 311f., 314
 - affirmative 32
 - empirische 51f.
 - negative 32
 - partikuläre 52
 - universelle 51
- Autonomer Prozess 282, 285f., 289ff.
- Axiom 73-78, 180, 272-278, 295-299, 345
- Axiomensystem 76, 78
- Axiome von Gulliksen 75, 79-83
- Axiome von Kolmogoroff 184
- Bedeutsamkeitsproblem 106
- Bedeutungskklärung 269f.
- Bedingter Reflex 20f, 23, 175f.
- Bedingungsdiskurs 68
- Bedingungszusammenhang 171, 239ff.
- Begründungsdiskurs 68
- Begründungsmuster 327, 329ff., 333
- Begründungszusammenhang 28, 30, 171
- Behauptung
 - analytische 54f.
 - a posteriori 55
 - a priori 55f.
 - empirische 55, 241
 - sachlogische 51, 241
 - synthetische 54ff.
- Behaviorismus 23f., 43-47, 63, 67, 177, 244, 293, 299, 312-315, 320
- Beobachtungsfehler 183, 197
- Beobachtungssprache 31-36, 233
- Beschreibungssprache 31-36
- Beschreibungstermini 231, 234, 271
- Bestätigungsgrad siehe Bewährungsgrad

- Bewährung (von Hypothesen) 143, 212
 Bewährungsgrad 143f., 157, 212, 219
 Beziehungsaspekt siehe Ausdrucksfunktion
 Binominalverteilung 203
 Blockdiagramm 287f.
- Chi-Quadrat 107, 121, 132, 210
 Command Aspect siehe Ausdrucksfunktion
 Conclusio 50, 164
- Darstellungsfunktion (von Sprechhandlungen) 274f., 277, 343f.
 Datenmatrix 120, 128, 131
 Deduktion 148f., 311
 Definiendum 41
 Definiens 41
 Definition 40f.
 - operationale 178, 254, 295
 Definitorischer Zirkel 52f, 64
 Deutung 34, 171f., 177, 227, 233, 239, 249ff., 257, 271, 318, 345
 Deutungsprädikator 33f., 231, 234f., 238, 241, 271
 Deutungssprache 237
 Deutungsterminus siehe Deutungsprädikator
 df siehe Freiheitsgrade
 Dichtefunktion 200, 203
 Differenzenskala 101, 105, 108, 111, 117
- Eigenname 33, 38f.
 Eigenverhalten 285f.
 Eindeutigkeitsproblem 103
 Einheitswissenschaft 63
 Ein-Klassen-Lösung (der LCA) 128f., 132
 Einsquantor 52, 271
 Elementaraussage 32f., 38, 41, 48, 56f., 232
 EM-Algorithmus 129f.
- Emotion 14, 24, 77, 152f., 227, 231ff., 270f.
 Empirischer Gehalt 87, 157, 297, 316f., 330
 Ereignis
 - sicheres 184, 186
 - unmögliches 163, 165, 174, 184
 - zufälliges 56, 185-189, 194, 198
 Erhebungsinstrumente
 - geschlossene 339
 - offene 258, 338f.
 Erkenntnisinteresse 63, 342
 Erklären siehe Erklärungsmodell
 Erklärung siehe Erklärungsmodell
 Erklärungsaufgabe 29, 66
 Erklärungsbegriff siehe Erklärungsmodell
 Erklärungsmodell 14f., 28, 47, 66ff., 87, 228, 241
 - deduktiv-nomologisches 68, 87-92, 141, 156f., 177, 183, 228, 239f., 241f., 246, 254ff., 305, 311
 - induktiv-statistisches 67f., 87, 91, 156, 182-220, 228, 239, 241, 257, 308, 326
 - intentionales 68, 87, 156, 171, 177, 234f., 260, 299, 305, 327
 - narratives 69, 87, 156, 254-258, 270, 300f., 327, 329-333
 Erkundungsexperiment 156
 Erwartungstreue 116, 191, 204
 Erwartungswert 80-83, 118, 130, 191f., 200-203, 207, 296
 Euklidische Geometrie 73, 78
 Euklidischer Raum 73, 88
 Ex-ante-Hypothesen siehe Hypothesenbildung ex ante
 Exemplarische Einführung (von Prädikatoren) 34
 Exhaustion 150
 Experiment 16ff., 35f., 55, 127-131, 141-179, 184f.,

- 188ff., 195, 198, 208f., 212f., 217-220
- Experimentelle Faktoren 144, 153
- Experimentelle Hypothesenprüfung siehe Experiment
- Explanandum 89, 91, 141, 183f., 255f.
- Explanans 255f.
- Faktorenanalyse 123
- Falsifikation 51, 143, 148, 157f., 163ff., 170, 207, 212, 240, 244-248, 297, 321, 346
- Fehlermodelle
 - mit fester Effektgröße 194ff., 219f.
 - mit zufälliger Effektgröße 196f., 220
- Feldexperiment 151, 153
- Feldstudie 151
- Fitnessbegrenzende Ressourcen 43, 315f.
- Formales System siehe Kalkül
- Forschung
 - explorative 259, 336
 - inspektive 258
- Forschungsmethoden
 - qualitative 15, 335-346
 - quantitative 15, 21
- Fragestellung, statistische
 - einseitige 210, 213-217, 262
 - zweiseitige 210, 213-217, 262
- Freiheitsgrade 121, 132
- Fremddeutung 249ff.
- Fremdverstehen 340f.
- Funktionalität (von Verhalten) 228-230, 322
- Gegensatz
 - kontradiktorischer 39f., 214f.
 - konträrer 39, 163, 234
 - polar-konträrer 40
- Gegenstandsverständnis (von Wissenschaft) 62-69, 239
- Gehalt
 - latenter 277
 - manifester 277
- Geltungsprüfung (von Hypothesen) 15, 28, 30, 65, 143, 156, 220, 262, 269, 340
- Generalisierung 148, 329
- Genetische Argumentation 246f., 269, 299ff.
- Genetische Rekonstruktion siehe genetische Argumentation
- Geschlossenes System 276, 285f.
- Gesetz der großen Zahlen 56, 187, 190-194, 206
- Gesetz des Effektes 22f., 311-315
- Gesetzesaussage
 - deterministische 183
 - statistische 182ff., 209
- Gesetzmäßigkeit
 - empirische 34, 68, 96, 98f., 142, 228, 232f., 241, 315
 - strukturelle 56, 73f., 78, 174, 305, 318, 320ff.
- Grundgesamtheit 201, 207, 217, 219
- Gruppendiskussion 338, 340
- Guttman-Skala 97-100, 103, 127
- Handlung 28, 42, 45ff., 53, 62, 66, 227f., 234-240, 245, 256, 260f., 318, 337
- Handlungsablauf 242f.
- Handlungsaspekt siehe Appellfunktion von Sprechhandlungen
- Handlungsergebnis 260f.
- Handlungsfolge 45, 47, 65, 67, 238, 241ff., 260f., 282, 285f., 290f., 293, 318, 320
- Handlungsgrund 68, 246ff.
- Handlungsorientierung
 - finale 263
 - finale 260f.
 - nonfinale 261ff.
- Handlungsprädikator 33f., 36, 238

- Handlungsregel 260f.
 Handlungszusammenhang 32,
 231, 238, 269, 271, 335
 Häufigkeit
 - absolute 191
 - relative 56, 184, 186f., 190ff.,
 206
 Häufigkeitsfunktion 199f.
 Häufigkeitsverteilung 107,
 199ff.
 Hermeneutische Analyse 16,
 254, 257ff.
 Hermeneutischer Zirkel 130,
 342ff.
 HO-Schema 89, 141, 311
 Homomorphie 94
 Hypothese
 - deterministische 184
 - probabilistische 184
 Hypothesenbildung ex ante
 149, 258, 339
 Hypothesengenerierung 336,
 339f.
 Hypothesenprüfung 141-178,
 211f., 257, 339f.
 Hypothetisches Konstrukt siehe
 theoretisches Konstrukt
- Idealisierung 91, 332
 Implikation 51
 Induktive Stützung 184, 308
 Informationsfunktion 205
 Inhaltsanalyse 340, 342
 Inhaltsaspekt siehe Darstel-
 lungsfunktion
 Inhaltsgleichheit (von Prädikato-
 ren oder Aussagen) 39
 Inspektion (der experimentellen
 Daten) 148
 Intention (Absicht) 24, 44, 47f.,
 68, 87, 155f., 171, 177,
 228, 230, 233ff., 238-
 252, 264, 277, 281, 291,
 299, 305, 319, 321f.,
 327f., 331
 Intentionalität siehe Intention
 Interaktion 153, 172, 257,
 272f., 278, 281f., 286,
 322f., 333, 341
- Interesse
 - im engeren Sinne 261ff., 271
 - im weiteren Sinne 261ff., 268,
 271, 282
 Internalisierte Intersubjektiv-
 tät 345
 Interpretation 83, 118, 168,
 171, 227, 263, 291, 335,
 340, 345
 Interpretationsfolie 263
 Interpretationsoffenheit 243ff.,
 277
 Interpunktion 273, 276
 Intervallschätzung 206
 Intervallskala 101f, 104-109,
 118
 Irreflexivität 36, 96, 103
 Itemcharakteristik 109-115,
 119, 199, 203
 Itemparameter 114-120
 Itemschwierigkeit 115, 119
 Item-Response-Modell 112
- Junktoren 48-51, 271
- Kalkül 49, 77f., 103, 286
 Kategorienwahrscheinlichkei-
 ten 128f., 131
 Kennzeichnung 33
 Klassische Testtheorie 74f.,
 78ff., 83, 112, 195, 198
 Kognitive Wende 24, 244
 Kommunikation 42, 249, 272-
 278, 282, 331, 337, 342
 - analoge 273, 277
 - digitale 273, 277
 - nichtsprachliche (nonverbale)
 273, 277, 341
 - parasprachliche 273, 277, 341
 Kommunikationstheorie 272
 Konditionierung
 - klassische 21f., 35
 - operante 22f.
 - rückwirkende 35
 - simultane 35
 - verzögerte 35, 40
 Konfidenzgrenzen 207
 Konfidenzintervall 206f.

- Konfidenzzahl 207
 Konfundierung 150
 Konjunktion 48f., 51, 271
 Konsistenz (von Bedeutungen) 344f.
 Konsistenz (von Schätzfunktionen) 206
 Konstellation (handlungsauslösende) 242-247, 263, 281f., 285f., 290, 343
 Kontext
 - externer 275, 343f.
 - interner 275, 343f.
 Kontextabhängigkeit (von Bedeutungen) 254, 269-272, 275
 Kontingenz 54, 160, 170-174, 184, 186f.
 Kontingenztafel 107
 Kontrollbedingung 154, 208
 Kontrollgruppe 153f., 208
 Kopula 32
 Korrelation 75, 82f., 107f., 123, 199-202
 Kovarianz 201f.
 Kritisierbarkeit 160, 163-167, 249, 251f.
 Kulturelle Regel 278, 335f.
- Laborexperiment 151, 153f.
 Latent-Class-Modell 10, 112, 121, 127-132, 194, 198, 233
 Latent Structure Analysis 113
 Latent-Trait-Modell 10, 112ff., 121f., 194, 198
 Lebensorientierung siehe Interessen im weiteren Sinne
 Likelihood 114, 130ff., 198, 204
 - bedingte (conditional) 115f., 120, 205
 - marginale 128, 130-132, 199
 Lineares Modell 107
 Lineares logistisches Modell 109, 194
 Logik 16, 30, 48-51, 54, 57, 317
 Logische Partikel 48-52
 Logistische Funktion 111, 114, 118
 Logistisches Testmodell siehe Rasch-Modell
 Maschine siehe Turing-Maschine
 - nichttriviale 283ff.
 - triviale 283ff., 293
 Maßzahltransformation 103f., 107ff., 111
 Maximum-Likelihood Methode (ML-Methode) 114, 198, 203f., 206
 - bedingte (CML-Methode) 115
 - gewichtete (WML-Methode) 115f.
 - marginale (MML-Methode) 115
 Membership-Wahrscheinlichkeit 130, 132
 Menschenbild 63
 Messung per fiat 37
 Messfehler 75f, 79-83
 Messinstrument 96, 103, 118, 156
 Messtheorie
 - allgemeine 74, 102-109
 - angewandte 102-109
 Metakommunikation 272-278
 Methoden
 - erfahrungswissenschaftliche 16, 159, 258
 - qualitative 15, 258, 336
 - quantitative 15, 258
 Mittel 44, 65f., 68, 171, 230f., 238-244, 265, 271, 291, 299, 322f., 330f.
 Mittelwert 75, 80, 107f., 199, 201f.
 Modell (eine Kalküls) 78, 80, 185-190, 286
 Modellgeltungstests 118, 121
 Modus ponens 50f., 89, 148
 Modus tollens 50f., 148
 Momente einer Verteilung 202
 Momentenmethode 203
- Narratives Interview 338
 Naturgesetz 19, 67, 74, 87f., 91, 143, 157, 182, 305

- Nebenfolgen siehe Folgen
 Negation 48f., 160, 171, 214, 271
 Neobehaviorismus 299, 293
 Newton'sche Mechanik 67, 86, 88, 183
 Nominalskala 95, 100f., 104, 107, 110, 141
 Nominator 32f., 38
 Normalverteilung 109, 113, 116, 203, 207, 296, 298
 Nullhypothese 120, 132, 209-216, 262

 Objektseitig definiert 15, 19f., 24, 47f., 96f., 108, 171, 175, 227-236, 242, 244, 263, 271f., 283, 291, 320, 327
 Operationalisierung 20, 36f., 67, 112, 116, 122, 160, 175-178, 295, 320f., 339
 Operationalisierbarkeit 160, 170, 175-178, 293, 313
 Ordinalskala 37, 99f., 103f., 107, 113, 116
 Ordnungsrelation 36, 96ff., 102
 - mehrstellige 102
 - zweistellige 102f., 113

 Parade Aspect siehe Appellfunktion
 Parameter
 - explizite 202ff.
 - implizite 201ff.
 Parameterschätzung 114, 117, 119f., 199, 201-207
 Partnerzentriertes Gespräch 338
 Personenparameter 114-117, 203
 Personenpopulation 81f., 198, 217, 219f.
 Perspektivendivergenz 45, 268, 276f.
 Poissonverteilung 203
 Prädikator
 - einstelliger 36
 - mehrstelliger 36f., 64
 Prädikatorenregel 37-41, 50, 162f., 231f.
 Pragmatische Evaluation 248-251, 299, 304
 Prinzip der Fremdheit 337f.
 Prinzip der Offenheit 337f.
 Prinzip der Reflexion 337f.
 Prinzip der Kommunikation 337f.
 Probabilistische Testtheorie
 siehe stochastische Testtheorie
 Produktmoment 202
 Produktmomentkorrelation
 siehe Korrelation
 Prototheorie 78, 96
 Prozessanalyse 326-333
 Prüfexperiment 156, 339
 Pseudoempirie 55, 160, 174, 316f.
 Psychoanalyse 26f., 43, 48, 238f., 249-252
 Psychologie (Definition) 14f.
 - experimentelle 20
 - geisteswissenschaftliche 63
 - naturwissenschaftliche 24, 28, 63, 67
 Psycho-Logik 74

 Quantitative Begriffe 36f.
 Quantor 51, 271
 Quotenstichprobe 218f.

 Randbedingung 88-91, 141, 143, 148, 183f., 229, 240, 242, 305
 Randomisierung 219
 Rangordnung 96, 98, 100f., 107, 152
 Rangskala siehe Ordinalskala
 Rasch-Modell 101, 113-120, 199
 Rationalskala 100f., 103, 105, 108
 Reaktion
 - bedingte (konditionierte) 21f., 35, 176
 - unkonditionierte 21f., 34, 175

- Reduktion
 - empirische 248-251, 269, 299, 302, 304
 - materielle 228f.
 - strukturelle 228f.f, 212
- Reflexionsterminus 231, 255f., 260, 263, 268, 335
- Regelkreis 156, 288f.
- Regelpfeil 38
- Regression 82f.
- Rekonstruktion (von Bedeutungen) 270f., 335, 343
- Relator siehe mehrstelliger Prädikator
- Reliabilität 74f.
- Report aspect siehe Darstellungsfunktion
- Repräsentationsproblem 102
- Repräsentativität 169, 216-220
- Responsivität (von Handlungen) 245
- Rollenspiel 338, 340
- Rückkoppelung 62, 283, 286, 289f.

- Sachhypothese 157, 159f., 164, 170-173, 211f., 215, 219
- Saturiertes Modell (der LCA) 128-132
- Schätzung unbekannter Parameter siehe Parameterschätzung
- Schätzfunktion 114, 117, 191, 201-207
- Selbstauskünfte 155, 176, 249, 270
- Selbstbeobachtung 20, 24, 241, 271
- Selbstdeutung 239, 249ff.
- Selbstmanifestationsaspekt siehe Ausdrucksfunktion
- Selbstreferentialität 276, 282-287
- Signifikanzniveau 121, 210f., 216
- Sinngehalt 171ff., 227f., 238f., 242-249, 254f., 294, 316
- Sinnkonstitution 341
- Sinnkrise 263f.
- Sinnstrukturierung 335
- Sinnzusammenhang 257, 261
- Situation 23f., 88, 152f., 156, 171, 177, 233, 242f., 245, 251, 254, 261, 263, 275, 282, 290f., 311f., 314f., 327, 335, 343
- Situationseinschätzung 239, 247
- Situationsveränderung
 - disfunktionale 315
 - funktionale 315
- Skala 37, 40, 102-112
- Skalierungsproblem 103, 114
- Skaleneigenschaften 100-112
- Skalentyp 100-112
- Sozialforschung
 - qualitative 130, 257f., 336-341, 345
 - quantitative 338f.
- Sprachnormierung 268
- Sprechhandlung 271-279
- Standardabweichung 107f., 118, 199-203, 207
- Statistik 115, 201-204
- Statistische Inferenz 10, 199-217
- Statistische Verfahren 123
 - nonparametrische 107, 109
 - parametrische 107f.
- Stichprobe 119, 122f., 153, 156, 169, 198ff., 202, 205, 217-219, 262, 303
 - abhängige 153
 - anfallende 217f., 220
 - geschichtete 218f.
 - unabhängige 153, 208f.
 - repräsentative siehe Repräsentativität
 - zufällige siehe Zufallsstichprobe
- Stichprobenauswahl 338
- Stichprobenfunktion 201f., 205
- Stichprobenmoment 203
- Stichprobenvarianz 199, 201
- Stochastische Modelle 194ff., 219
- Stochastische Testtheorie 112f.
- Störvariable (SV) 83, 150f., 217ff., 335
- Strenge Theorienprüfung 157ff.

- Streuung 75f., 205
- Strukturmerkmale 65, 154f.
- Subjektseitig definiert 15, 19f., 24, 47, 96f., 108, 171, 175ff., 227-236, 242ff., 255, 263, 271f., 282f., 320
- Subjunktion 48ff., 89, 143, 164f., 172, 240, 271, 287
- Succedens 311, 314
- Suffizienz (von Schätzfunktionen) 205
- Systematischer Meßfehler 79
- Systematischer Varianzanteil 75
- Systemtheorie 285f.
- Szientismus 86f.

- Tautologie 50, 164, 296, 311
- Technische Verwertbarkeit (der Naturwissenschaften) 28, 63ff.
- Teilnehmende Beobachtung 338, 341
- Temporalstruktur (des Handelns) 256
- Terminologiebildung 30
- Terminologische Regel 37-42, 50, 56
- Testanalyse 121
- Testkonstruktion 121, 304
- Text 275, 341f.
- Textinterpretation 341-345
- Theoretical Sampling 338
- Theoretisches Konstrukt 33-36, 238, 241, 249
- Theorieanteile
 - empirische 57, 297, 311-324
 - strukturelle 57, 74, 297, 311-324, 326
- Theoriesprache 31-36, 233
- Tierversuch 176f.
- Transitivität 36, 96, 103
- True-Score 75, 79-83, 198
- Turing-Maschine 283
- Typikalität 338

- Umweltbedingung 230, 243-247, 263, 283, 291, 294, 327
- Ungleichung von Rao-Cramér 205
- Ungleichung von Tschebescheff 192f.
- Ununterscheidbarkeit 186-189, 262
- Ursache-Wirkungs-Relation 19, 66ff., 239, 254, 287, 299

- Validität
 - externe 151, 154
 - interne 151, 220
- Variable
 - abhängige (AV) 141-144, 160, 171ff., 240, 295
 - binäre 95
 - dichotome 95, 110
 - latente 108-113, 121, 123, 127f.
 - manifeste 108-113, 123
 - polytome 95
 - qualitative 94-97, 109, 121, 127f.
 - quantitative 94-97
 - unabhängige (UV) 141-144, 160, 171ff., 240, 295
- Variablenstufen 143f., 153
- Variablenpsychologie 159, 321
- Varianz 75, 80f., 191f., 200ff., 205
- Varianzanalyse 195, 197
- Verdoppelung der Realität 13f., 165f., 170, 294f.
- Verhältnisskala siehe Rationalskala
- Verhalten
 - bloßes 227, 234f.
 - funktionales 43, 47, 228, 234f., 316
 - nonverbales 273
 - paraverbales 273
- Verifikation 54, 143, 246
- Vermögenspsychologie 13f., 126, 165
- Verstehen
 - interpersonales 239, 343, 345

- methodisch kontrolliertes 270f.
- Versuchsbedingung (VB) 144, 153, 155
- Versuchsgruppe 153f., 208
- Versuchsplan 146, 208
- einfaktorierter 144, 153
- mehrfaktorierter 144, 153
- multivariater 153
- univariater 153
- Verteilungseigenschaften 109
- Verteilungsfunktion 113, 188, 200, 202, 296
- Verwertungszusammenhang (von Wissenschaft) 63
- Vollständige Induktion 51f.

- Wahrheit
 - analytische 54ff.
 - a posteriorische 55
 - a priorische 55ff.
 - empirische 57, 245
 - (formal) logische 57
 - sachlogische 57
 - synthetische 54, 56f.
- Wahrheitsfähigkeit 28
- Wahrheitstafel 49ff, 162f., 165, 174
- Wahrheitswert 48, 106
- Wahrnehmungsprädikator 232
- Wahrscheinlichkeit 56, 102, 110, 114f., 128f., 168, 183-217, 320
- Wahrscheinlichkeitsdichte 115, 188, 192, 199f., 202f., 206
- Wahrscheinlichkeitsfunktion 190, 200
- Wahrscheinlichkeitsverteilung 80, 114, 199-203, 207, 211, 213
- Wertdefinitheit 48f.
- Wertfreiheit 63, 159, 268
- Widerfahrnis 234f.
- Widerspruchsfreiheit 41, 50, 76, 160-163
- Widerspruchshaltigkeit 164
- Wirksamkeit (von Schätzfunktionen) 205
- Wissenschaftssprache 15, 32ff., 37, 41, 77, 255, 271, 277
- Wissenschaftstheorie 16, 54, 56, 78, 86
- Wissensideal 62-68, 239
- aristotelisches 73-84
- galileisches 66f, 86f.
- Würfel
 - regelmäßiger 185f., 262
 - unregelmäßiger 189

- Ziel-Mittel-Relation 44, 66, 68, 235-241, 330
- Zirkelschluss 13, 164f.
- Zirkularität 166, 311-317
- Zufall 184-220, 271
- Zufälligkeit siehe Zufall
- Zufallsauswahl 195f., 217-220, 338
- Zufallsexperiment 184-190, 194f., 198, 217-220
- Zufallsfehler 80, 83, 119f., 195f.
- Zufallsgenerator 56, 92, 262
- diskreter 185-189
- kontinuierlicher 185, 187ff.
- Zufallsprozess siehe Zufallsexperiment
- Zufallsstichprobe 199, 217f.
- Zufallsvariable 81, 115, 192, 195, 197, 202ff., 206
- diskrete 192
- stetige 192, 204
- Zulässige Maßzahltransformation 103f., 111ff., 115
- Zweck 230, 235f., 238, 260, 271, 299
- Zweckmäßigkeit 47, 77, 79, 320, 326

Wilhelm Kempf**Forschungsmethoden der Psychologie**
Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher
Hermeneutik
Band II: Quantität und Qualität

(Hochschullehrbücher, Bd. 3). 2008. 396 S., brosch., mit 47 Abbildungen und 59 Tabellen,
€ 44,90. ISBN 978-3-936014-15-0.

Gegenstand des zweiten Bandes der Forschungsmethoden der Psychologie sind zwei Teilgebiete der Methodenlehre, die auf geradezu entgegengesetzten Polen des Methodenspektrums angesiedelt sind: (1) Textinterpretation und Inhaltsanalyse, die dem Wissensideal der qualitativen Sozialforschung verpflichtet sind, und (2) Psychometrie, die von manifesten Beobachtungsdaten auf zugrunde liegende Eigenschaften der Probanden schließt und die wohl konsequenteste Umsetzung eines quantifizierenden Wissensideals in der Psychologie darstellt.

In didaktischer Hinsicht dem Prinzip des Lernens am konkreten Objekt verpflichtet, will das Buch den Leser in die Lage versetzen, die einschlägige Fachliteratur kritisch zu rezipieren, die Methoden in der eigenen Forschungspraxis anzuwenden und sie in gegenstandsangemessener Weise miteinander zu verbinden. Nicht nur die Anwendungsnähe der Methoden, sondern auch ihre theoretische und (im Falle der Psychometrie) mathematische Grundlegung stehen daher im Zentrum des Buches.

Aus dem Inhalt

1. Textinterpretation und Inhaltsanalyse: Qualitative Sozialforschung / Grundlagen der sozialpsychologischen Rekonstruktion / Textinterpretation / Qualitative und quantitative Inhaltsanalyse / Integration quantitativer und qualitativer Analysemethoden
2. Psychometrie: Grundlagen / Testen und Messen / Klassische und stochastische Testtheorie / Parameterschätzung, Modellkontrolle und Beurteilung der Testleistung der Probanden / Rasch-Modell / Verallgemeinerungen des Rasch-Modells / Latent-Class-Analyse / Modellvergleiche / Validität

Wilhelm Kempf & Markus Kiefer (Hrsg.)
Forschungsmethoden der Psychologie
Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher
Hermeneutik
Band III: Natur und Kultur

(Hochschullehrbücher, Bd. 4). 2009. 366 S., brosch., mit 63 Abbildungen und 8 Tabellen,
€ 44,50. ISBN 978-3-936014-17-4.

Der dritte Band der Forschungsmethoden der Psychologie beschreibt auf verständliche und anschauliche Weise wichtige Methoden der natur- und kulturwissenschaftlichen Psychologie und arbeitet deren spezifischen Stellenwert für das gesamte Fachgebiet heraus. Er führt in die wichtigsten Verfahren der qualitativen Sozialforschung zur Rekonstruktion der subjektiven Wirklichkeit des Menschen ein, beschreibt die Methoden der behavioralen experimentellen Psychologie und gibt einen Überblick über neurowissenschaftliche Methoden in der Psychologie.

Das Lehrbuch plädiert für den gegenstandbezogenen Einsatz verschiedenster Forschungsmethoden, der die klassischen Grenzen zwischen Natur- und Kulturwissenschaften aufhebt und so der Komplexität des Forschungsgebiets "Mensch" gerecht wird. Es bleibt aber nicht bei einer Einführung in das Methodenspektrum der verschiedenen psychologischen Strömungen – entworfen wird die Skizze einer integrativen Psychologie, in der die dem Gegenstand der Psychologie inhärente Dialektik von Subjektivität und Objektivität zur Synthese gebracht wird.

Das Buch gibt nicht nur Studierenden und Doktoranden unerlässliches theoretisches und praktisches Rüstzeug für die Durchführung ihrer Forschungsarbeiten in die Hand, sondern wendet sich darüber hinaus auch an alle diejenigen, die sich für die wissenschaftstheoretische und methodologische Entwicklung der akademischen Psychologie interessieren.

Aus dem Inhalt: 1. Psychologie als Natur- und Kulturwissenschaft (Wilhelm Kempf) 2. Die soziale Konstruktion der Wirklichkeit (Wilhelm Kempf) 3. Methodologie und Methodik der Grounded Theory (Günter Mey & Katja Mruck) 4. Behaviorale Methoden der experimentellen Psychologie (Lynn Huestegge & Iring Koch) 5. Neurowissenschaftliche Methoden (Markus Kiefer) 6. Die Konvergenz der Paradigmen (Klaus Hönig)