

*Vorlesung (HS-2011):
Wissenschaft und
Pseudowissenschaft in der
Psychologie*

Autor: Siegfried Macho

Text redigiert bis Seite 124

Datum der letzten Änderung: Mittwoch, 30. November 2011

Inhaltsverzeichnis

1. Einleitung: Das Problem der Unterscheidung von Wissenschaft und Pseudowissenschaft	1
1.1 Übungen zu Kapitel 1	8
2. Konzeptionen der Psychologie als Wissenschaft	9
2.1 Die induktiv-empiristische Konzeption wissenschaftlicher Tätigkeit	9
2.1.1 Exkurs: Verschiedene Arten von Schlussfolgerungen	9
2.1.2 Prinzipien der empiristisch-induktiven Charakterisierung der Wissenschaften	15
2.1.3 Exkurs: Zu Bedeutung und Missbrauch der Faktorenanalyse in der psychologischen Forschung	18
2.2 Die deduktiv-falsifikationistische Konzeption der Wissenschaft	22
2.2.1 Poppers Kritik an der induktiv-empiristischen Konzeption	22
2.2.1.1 Poppers Kritik der Induktion	22
2.2.1.2 Poppers Kritik an der Trennung von Beobachtungs- und theoretischer Sprache	25
2.2.2 Poppers kritischer Rationalismus	26
2.2.3 Poppers deduktiv-falsifikationistische Konzeption der Wissenschaft	29
2.2.3.1 Experimentum Crucis	50
2.3 Konzeptionen der Wissenschaft nach Popper	57
2.3.1 Zwei grundlegende Probleme: Holismus und die Unterbestimmtheit von Theorien	57
2.3.1.1 Das Problem des Holismus	58
2.3.1.2 Die Unterbestimmtheit von Theorien durch Daten	60
2.3.1.2.1 Die Theorie des Schwarzen Kastens	60
2.3.1.3 Strategien im Umgang mit Holismus und Unterbestimmtheit von Theorien durch Daten	68
2.3.2 Die Logik der Forschungsprogramme von Lakatos und die Idee des Paradigmenwechsels von Kuhn	70
2.3.2.1 Die Logik der Forschungsprogramme nach Lakatos	70
2.3.2.2 Die Idee des Paradigmenwechsels nach Kuhn	84
2.3.3 Die strukturalistische Konzeption von Wissenschaft	88
2.3.3.1 Die Menge der partiellen potentiellen Modelle einer Theorie	93
2.3.3.2 Die Menge der potentiellen Modelle einer Theorie	96

2.3.3.3	Die Menge der Modelle einer Theorie	99
2.3.3.4	Beschränkungen in partiellen Modellen und Verbindungen zwischen partiellen Modellen	100
2.3.3.5	Intendierte Anwendungen	102
2.3.3.6	Würdigung: Stärken und Schwächen des Ansatzes	102
2.4	Übungen zu Kapitel 2	104
3.	Wissenschaftliche Erklärungen in der Psychologie	107
3.1	Die Struktur wissenschaftlicher Erklärungen	107
3.1.1	Die »Grossmutter« der Struktur wissenschaftlicher Erklärungen: Das Hempel-Oppenheim-Schema	107
3.1.1.1	Das Hempel-Oppenheim-Schema der Erklärung: Deduktiv- nomologische Form	108
3.1.1.2	Das Hempel-Oppenheim-Schema der Erklärung: Induktiv- probabilistische Version	109
3.1.2	Einwände gegen das Hempel-Oppenheim-Schema	113
3.1.2.1	Die Erfüllung des HO-Schema ist nicht hinreichend	113
3.1.2.2	Die Erfüllung des HO-Schema ist nicht notwendig	115
3.1.3	Erklärung als Aufdeckung des kausalen Mechanismus	116
3.1.4	Funktionale Erklärungen in der Psychologie	117
3.1.5	Hermeneutik und kausale Erklärungen	120
3.1.6	Probabilistische Erklärungen in der Psychologie	125
3.2	Ebenen der Erklärung eines Informationsverarbeitungssystems	125
3.3	Psychologische Erklärungen und Reduktionismus	125
3.4	Scheinerklärungen in der Psychologie	125
4.	Theorien und Theoriendynamik in der Psychologie	126
4.1	Der Einfluss der Konzeption von Wissenschaft auf die psychologische Forschungspraxis	126
4.2	Theorienbildung in der Psychologie	126
4.3	Probleme der Prüfung von Theorien in der Psychologie	126
4.3.1.1	Klassische Signifikanztests und die Prüfung von Theorien	126
5.	Pseudowissenschaftliche Praktiken in der Psychologie	127
6.	Verwendete Literatur	128

1. Einleitung: Das Problem der Unterscheidung von Wissenschaft und Pseudowissenschaft

Betrachtet man die Geschichte verschiedener naturwissenschaftlicher Disziplinen, so stösst man auf pseudowissenschaftliche Theorien und Praktiken, welche eine Art eigenständige Existenz führen, parallel zu den eigentlichen Disziplinen. Typische Beispiele hierfür sind:

- Alchemie (Chemie),
- Astrologie (Astronomie),
- Kreationismus (Darwinismus / Biologie),
- Quacksalberei (Medizin).

Interessanterweise findet man immer wieder Fälle von prominenten Wissenschaftlern, die sich neben ihrer wissenschaftlichen Tätigkeit ausführlich mit pseudowissenschaftlichen Theorien und Praktiken beschäftigten. Bekannte Beispiele hierfür sind Isaac Newton (1642–1727), der vermutlich mehr Zeit seinen alchemistischen Studien als seiner wissenschaftlichen Tätigkeit widmete (Westfall, 1980), oder Johannes Kepler (1571–1630), der für Wallenstein Horoskope erstellt haben soll.

Die Psychologie als naturwissenschaftliche Disziplin bildet neben der Medizin einen reichhaltigen Nährboden für pseudowissenschaftliche Theorien und Praktiken. Pseudowissenschaftliche Tätigkeiten im Rahmen der Psychologie reichen von Aktivitäten bestimmter Sekten (z.B. Dianetik der Scientologen) über New Age Therapien (z.B. Rebirthing) bis zu zweifelhaften diagnostischen Methoden (z.B. Graphologie).



Bemerkung

Die Psychologie als wissenschaftliche Disziplin wird als ein Zweig der Naturwissenschaften betrachtet, da ihre Theorien dazu dienen, Phänomene der realen Welt zu erklären und vorherzusagen und daher einer empirischen Prüfung unterzogen werden können und müssen.

Demgegenüber machen geisteswissenschaftliche Disziplinen, wie Philosophie und reine Mathematik, keine empirisch prüfbareren Aussagen und können daher im »Lehnstuhl« betrieben werden (was keineswegs bedeutet, dass sie einfacher sind).

Nun stellt sich sofort die folgende Frage:



Frage:

Wie lässt sich Wissenschaft von Pseudowissenschaft unterscheiden?

Die Beantwortung dieser Frage ist nicht nur von theoretischer sondern auch von eminent praktischer Bedeutung. So versuchen in der USA immer wieder Vertreter des so genannten »Intelligent Designs« – einer

Weiterentwicklung des Kreationismus (siehe Kitcher, 2009) – juristisch eine Anerkennung bzw. eine Gleichberechtigung dieser Lehre mit der darwinistischen Evolutionstheorie durchzusetzen. Um derartige Versuche abzuwehren, muss gezeigt werden, dass es sich bei der Lehre vom »Intelligent Design« um eine Pseudowissenschaft handelt.

Nun gibt es eine Reihe von Indikatoren für das Vorliegen von pseudowissenschaftlichen Theorien und/oder Praktiken. Methode 1-1 fasst die wichtigsten zusammen.



Methode 1-1: Indikatoren pseudowissenschaftlicher Praktiken (Lawson, 2007):

1. Ungenaue, wissenschaftlich klingende Sprache:

Verwendung von wissenschaftlich klingenden Ausdrücken wie »holistisch«, »Energie«, »Rhythmen«, »Strahlen«, »Felder«, etc.

2. Hinweis auf zweifelhafte, idiosynkratische Praktiken:

Verfahren sind nur wirksam, wenn sie in ganz bestimmter Weise durchgeführt werden, z.B. Medikamente, welche nur dann wirksam sind, wenn sie auf bestimmte Weise geschüttelt werden.

Meist sind nur der/die »Eingeweihte« in der Lage, diese Prozeduren korrekt auszuführen.

3. Subjektive persönliche Erfahrungen als hauptsächliche Evidenz:

Wirkungen von Verfahren werden nicht durch kontrollierte Studien ermittelt, sondern basieren auf Zeugnissen von Einzelpersonen.

Diese sind oft als Bekenntnis- bzw. Empfehlungsschreiben den Werbungen für ein Medikament oder eine Methode beigelegt.

4. Vorspiegelung des Vorhandenseins fundierter empirischer Evidenz:

Oft wird auf eine lange Liste wissenschaftlicher Publikationen verwiesen, die in den meisten Fällen in irgendwelchen fragwürdigen Zeitschriften erschienen sind. Somit wird der Eindruck des Vorliegens von konvergierender bzw. kumulativer Evidenz erweckt.

Bemerkung:

Es gibt allerdings auch Fälle, wo extrem renommierte Zeitschriften derartig fragwürdige Ergebnisse publizierten, z.B.: *Nature*, *Psychological Bulletin* und *Journal of Personality and Social Psychology*.

5. Hinweis auf zweifelhafte Autoritäten:

Oft werden irgendwelche Autoritäten zitiert, die niemals auf dem Gebiet wissenschaftlich tätig waren bzw. keinerlei fundiertes Wissen auf dem Gebiet besitzen.

Bsp.: »Einstein sagt: Wir verwenden nur 10% unseres Gehirns.«

Bemerkung:

Niemand weiss, ob Einstein je etwas in diese Richtung behauptet hat (Die Forschungsassistenten des Albert Einstein-Archivs konnte keinerlei Hinweise finden [Beyerstein, 2007, Seite 54]).

6. Hinweis auf alte – scheinbar verloren gegangene – Weisheiten:

Oft wird darauf verwiesen, dass es sich bei einer Methode um eine alte Weisheit der Menschen handelt, die aber im Laufe der Jahrhunderte verloren ging oder von irgendwelchen Urvölkern überliefert wurde.

Bsp.: Weisheit der alten Chinesen oder Wissen des Indianerhäuptlings »Sitting Bull«.

7. Aussagen über Zusammenhänge, die in krassem Widerspruch zu unserem physikalischem Grundwissen stehen:

Es werden Aussagen gemacht, die in krassen Gegensatz zu dem uns bekannten physikalischen Wissen stehen.

Bsp.: Wasser übernimmt die Eigenschaften jedes Stoffes, welcher sich irgendwann einmal darin befunden hat. Daher wirkt der Stoff weiter, auch wenn er nicht mehr mit physikalisch-chemischen Mitteln nachweisbar ist.

8. Umkehrung der Beweislast:

Anstelle des Erbringens eines Beweises, dass eine Entität oder eine Wirkung vorhanden ist, wird – umgekehrt – vom Gegenüber gefordert, den Beweis zu erbringen, dass besagte Entität oder Wirkung nicht besteht.

Bsp.: Vertreter von Religionen weisen darauf hin, dass man die Nichtexistenz Gottes nicht beweisen kann.

9. Rückzugsgefechte anstelle von Fortschritt:

Gute wissenschaftliche Theorien zeichnen sich dadurch aus, dass Sie zu neuen empirische prüfbareren Vorhersagen und zu neuen Problemstellungen führen.

Pseudowissenschaftliche Theorien machen keine prüfbareren Vorhersagen. Stattdessen führen Sie ständig Rückzugsgefechte, in Form von Neuinterpretationen oder der teilweisen Aufgabe von Positionen (ohne Aufgabe der Kernaussagen).

Bsp.: Die Entwicklung des Kreationismus zum »Intelligent Design« ist ein typisches Beispiel: Hier wurden alte Positionen aufgegeben (nämlich, dass Gott die Erde vor ungefähr 6000 Jahre erschaffen hat. Die zentrale Position, dass der Mensch (zumindest teilweise) von Gott geschaffen wurde, wird jedoch beibehalten. Allerdings wird dies verschleiert, indem von einer »Intelligenz« gesprochen wird, die teilweise in die Evolution eingreift (Kitcher, 2009).

Die in Methode 1-1 aufgeführten Punkte sprechen für sich und es lassen sich zahlreiche Beispiele von pseudowissenschaftlichen Praktiken aus Medizin und Psychologie finden, auf welche eine oder mehrere der angeführten Indikatoren zutreffen (Übung 1-1).

Auf der anderen Seite gab es zahlreiche Versuche, Kriterien für Wissenschaftlichkeit aufzustellen. Methode 1-2 präsentiert eine Auswahl.



Methode 1-2: Eine Auswahl möglicher Kriterien zur Identifikation von Wissenschaft:

Es wurde eine Anzahl möglicher Merkmale vorgeschlagen, durch welche sich die Wissenschaft angeblich auszeichnet. Hier eine Auswahl:

1. Wissenschaft verwendet klar definierte und unumstrittene Konzepte.
2. Bei den von den Wissenschaften postulierten bzw. entdeckten theoretischen Entitäten wie Kräfte, Felder, Repräsentationen, kognitiven Prozessen, etc. handelt es sich zweifelsfrei um real existierende Dinge, welche den beobachtbaren Erscheinungen zugrunde liegen.
3. Die meisten Modelle der exakten Naturwissenschaft bilden die Wirklichkeit korrekt ab.
4. Wissenschaft nähert sich immer mehr der Wahrheit an
5. Theorien, die durch empirische Daten falsifiziert wurden, werden aus dem Kanon der wissenschaftlichen Theorien ausgeschlossen.
6. Wissenschaft basiert auf der Anwendung fundierter Methoden, die sich als zuverlässig erwiesen haben.

Einige der in Methode 1-2 angeführten Punkte kann man in Gesprächen mit Laien über die Natur der Wissenschaft wieder finden.

Kommen wir zurück zur oben gestellten Frage bezüglich der Abgrenzung von Wissenschaft und Pseudowissenschaft. Die beiden Methoden scheinen uns der Lösung dieses Problems näher zu bringen, da sie uns scheinbar klare Kriterien liefern, wie sich wissenschaftliche und pseudowissenschaftliche Praktiken identifizieren lassen. Leider ist diese Vorgehensweise aus folgenden Gründen nicht Ziel führend:

- Keines der in Methode 1-2 aufgeführten Kriterien dient zur Identifikation wissenschaftlicher Tätigkeit. In der Tat *gibt es bis heute kein unumstrittenes Kriterium, welches es ermöglicht, wissenschaftliche Praxis von anderen Arten von Tätigkeiten zu trennen.*
- Es gibt Grenzbereiche bzw. Grauzonen, für welche ungeklärt ist, ob es sich um eine Wissenschaft handelt oder eher um eine Pseudo- oder Protowissenschaft (d.h. eine Disziplin, die noch nicht als vollwertige Wissenschaft anerkannt ist). Die Psychoanalyse bietet hierfür ein gutes Beispiel.

Die Tatsache, dass es keine stringenten Kriterien gibt, um Wissenschaft von anderen Tätigkeiten zu trennen, hat eine Reihe von Wissenschaftsphilosophen dazu verführt, eine Art Gleichsetzung zwischen Wissenschaft und Mythos zu postulieren (Quine 1951; Feyerabend, 1995; Hübner, 1985).



Bemerkung

Fairerweise sollte hinzugefügt werden, dass es zwischen den drei Autoren klare Unterschiede gibt hinsichtlich ihrer kritischen Einstellung zur Wissenschaft. So ist Quine sicherlich weniger wissenschaftskritisch (oder sogar –feindlich) eingestellt als die beiden anderen Autoren. Dies belegt der folgende Vergleich von Zitaten aus den Arbeiten von Quine und Feyerabend:

Physical objects are conceptually imported into the situation as convenient intermediaries – not by definition in terms of experience, but simply as irreducible posits comparable, epistemologically, to the gods of Homer. Let me interject that for my part I do, qua lay physicist, believe in physical objects and not in Homer's gods; and I consider it a scientific error to believe otherwise. But in point of epistemological footing the physical objects and the gods differ only in degree and not in kind. Both sorts of entities enter our conception only as cultural posits. The myth of physical objects is epistemologically superior to most in that it has proved more efficacious than other myths as a device for working a manageable structure into the flux of experience (Quine, 1951).

Bei Feyerabend (1995, Seite 385) liest man hingegen:

Es gibt also keinen klar formulierten Unterschied zwischen Mythen und wissenschaftlichen Theorien. Die Wissenschaft ist eine der vielen Lebensformen, die die Menschheit entwickelt haben, und nicht unbedingt die beste. Sie ist laut, frech, teuer und fällt auf. Grundsätzlich überlegen ist sie aber nur in den Augen derer, die bereits eine gewisse Position bezogen haben und die die Wissenschaft akzeptieren, ohne jemals ihre Vorzüge und Schwächen geprüft zu haben.

Kurt Hübner nimmt hier eine Mittelposition ein, wobei er im Laufe der Zeit immer mehr der Position von Paul Feyerabend zuneigte.

An einer radikalen Gleichsetzung zwischen Mythos und Wissenschaft – wie sie Feyerabend (1995) und auch Hübner (1985) sehen – kann etwas nicht stimmen, denn es ist offensichtlich, dass die wissenschaftlichen Vorhersagen hinsichtlich Genauigkeit und Detailliertheit jene aller mythologischen Systeme und auch jene des Alltagsverstandes bei weitem übertrifft. Des Weiteren führen die aus der wissenschaftlichen Forschung gewonnen Erkenntnisse zu technologischen Fortschritten, welche das Leben der Menschheit fundamental verändert hat [Vergleiche hierzu den letzten Satz im obigen Zitat von Quine (1951)].

Die Gleichsetzung von Wissenschaft mit Mythos, Pseudowissenschaft etc. birgt auch erhebliche Gefahren in sich. Sie kann nämlich dazu führen, dass gewisse wissenschaftliche Standards, welche für den Fortschritt einer Disziplin relevant sind, aufgeweicht werden. Das klassische Beispiel hierfür bildet die (evidenzbasierte) Medizin. Nicht nur bestehen neben Heilmitteln, deren Wirkung wissenschaftlich getestet wurde, zahlreiche Techniken und Heilmittel aus dem Bereich der Alternativ- (oder Komplementär-) Medizin, deren Wirksamkeit äußerst umstritten ist, sondern es gibt auch zahlreiche Versuche, die Kriterien für die Prüfung der Wirksamkeit von Therapien bzw. Medikamenten aufzuweichen. Genau auf diesen strengen Methoden beruht jedoch der Erfolg der evidenzbasierten Medizin (siehe hierzu Ernst & Singh, 2009). Wir werden später noch näher auf diese Problematik eingehen, welche auch die klinisch-psychologische Praxis betrifft.

Fassen wir das bisher Gesagte zusammen:

1. Es gibt kein eindeutiges Kriterium oder Verfahren, welche es erlauben würde, wissenschaftliche Theorien und Praktiken in eindeutiger Weise von anderen Tätigkeiten zu unterscheiden. Auch ist die Wissenschaft keineswegs gefeit gegen »pseudowissenschaftliche Anfechtungen«, was erklärt, warum auch in seriösen Zeitschriften gelegentlich zweifelhafte Arbeiten publiziert werden.
2. Auf der anderen Seite hat sich die Wissenschaft als die bisher erfolgreichste Unternehmung der Menschheitsgeschichte herausgestellt, Ereignisse zu erklären und vorherzusagen (und damit zu kontrollieren und manipulieren). Es ist fair zu behaupten, dass alle anderen Versuche – wie Alltagserklärung, Magie, Mythos, Religion, etc. – sich der Wissenschaft als deutlich unterlegen gezeigt haben.

Diese beiden Aussagen scheinen auf den ersten Blick leicht widersprüchlich, dann wenn es keine eindeutigen Kriterien, wissenschaftliche Tätigkeit von anderen zu sondern, warum ist letztere dann so erfolgreich?

Dies hat folgenden Grund: Wissenschaft kann zwar – wie die meisten Kategorien – als eine unscharfe Kategorie betrachtet werden. Sie zeichnet sich jedoch durch das vermehrte Auftreten von bestimmten »positiven« Eigenschaften auf. Hierbei sind vor allem zwei zu nennen:

1. Die Tendenz bzw. Forderung, alle Annahmen und Beweise auf den Tisch zu legen, so dass sie der allgemeinen Kritik der Forschergemeinschaft unterworfen werden können.
2. Eine kontinuierliche Kritik, von der nichts ausgeschlossen ist.

Diese beiden Aspekte zusammengenommen wurden von vielen als das definierende Merkmal von Wissenschaft betrachtet. Dies ist insofern nicht ganz korrekt, als die Leute auch im Alltag durchaus kritisch denken und handeln (oder es zumindest versuchen). Andererseits mahlen die »Mühlen der Kritik« auch in der Wissenschaft oft ziemlich langsam. Dies trifft – wie wir noch sehen werden – auch auf die psychologische Forschung zu.



Bemerkung zu den unscharfen Kategorien;

Wie oben angedeutet, kann die Wissenschaft als eine unscharfe Kategorie bezeichnet werden, die sich nicht durch klare Kriterien von anderen menschlichen Tätigkeiten abgrenzen lässt.

Die Idee unscharfer Kategorien wurde von Wittgenstein (1953) hervorgehoben. Er betrachtet verschiedene Arten von Spielen (z.B. Schach, oder ein Kind wirft einen Ball gegen eine Wand) und fragt, was das gemeinsame Merkmal all dieser Tätigkeiten sei. Er kommt zu dem Schluss, dass es keinen derartigen gemeinsamen Nenner gibt. Vielmehr zeichnen sich Kategorien durch so genannte Familienähnlichkeiten aus. Damit ist gemeint, dass die Mitglieder einer Kategorie untereinander ähnlich sind, was aber nicht ausschliesst, dass zwei verschiedene Exemplare an unterschiedlichen Polen der Kategorie sich sehr unähnlich sein können.

Diese Idee wurde auch innerhalb der Psychologie zur dominierenden Auffassung (siehe z.B. Rosch & Mervis, 1975).

Die folgende Darstellung beschäftigt sich daher mit den folgenden Fragen:



Fragen:

1. *Wie funktioniert Psychologie als Wissenschaft?*
2. *Wie funktionieren gängige pseudo- und protowissenschaftliche Praktiken und wie unterscheiden sich diese von den wissenschaftlichen Praktiken?*
3. *Welche Gründe gibt es für die grosse Verbreitung pseudowissenschaftlicher Praktiken in den unterschiedlichen Gebieten?*

Um diese Fragen zu beantworten, werde ich wie folgt verfahren:

- Ich beginne mit der Darstellung verschiedener Auffassungen von Wissenschaft und zeige auf, wie Psychologie in diese vorgeschlagenen Schemata passt.

- Das Problem der wissenschaftlichen Erklärung wird behandelt und es wird untersucht, was eine gute psychologische Erklärung ausmacht.
- Die Theoriendynamik, d.h. die Bildung und Verwerfung von psychologischen Theorien und Forschungsprogrammen wird untersucht.
- Wir betrachten verschiedene pseudowissenschaftlichen Praktiken und untersuchen, inwieweit Sie den in Methode 1-1 angeführten Kriterien entsprechen.
- Wir wenden uns der Frage nach zentralen psychologischen Mechanismen des Glaubens an den Erfolg pseudowissenschaftlicher Praktiken zu.

1.1 Übungen zu Kapitel 1



Übung 1-1:

Diskutiere eine dir bekannt pseudowissenschaftliche Methode / Theorie und anhand der gegebenen Indikatoren. Welche der Indikatoren sind auf diese Methode anwendbar und welche nicht?



Übung 1-2:

Diskutiere die in Methode 1-2 gegebenen Kriterien. Gib Beispiele für das Zutreffen oder Nichtzutreffen der verschiedenen Kriterien.

2. Konzeptionen der Psychologie als Wissenschaft

Die unterschiedenen Konzeptionen von Wissenschaft lassen sich in 3 grobe Kategorien aufteilen, wobei die Auffassung von Karl Popper (1902–1994), dem vermutlich wichtigsten Wissenschaftstheoretiker des 20. Jh., den Dreh- und Angelpunkt bildet:

- (a) Die vorpopperianische induktiv-empiristische Konzeption.
- (b) Die deduktiv-falsifikationistische Konzeption von Popper.
- (c) Postpopperianische Konzeptionen, die sich aus der Kritik an der Auffassung Poppers ergaben.

Im Folgenden wird untersucht, wie die wissenschaftlich orientierte Forschung in der Psychologie unter den einzelnen Konzeptionen aufgefasst wird.

2.1 Die induktiv-empiristische Konzeption wissenschaftlicher Tätigkeit

Diese Konzeption der Wissenschaft geht auf den englischen Philosophen Francis Bacon (1561–1626) zurück (Andere wichtige Vertreter sind David Hume (1711–1776) und John Stuart Mill (1806–1873)). Im 20. Jh. wurde diese Position in moderner Fassung von den Mitgliedern des *Wiener Kreises* vertreten.

Bevor wir uns den zentralen Postulaten dieses Ansatzes zuwenden, muss eine Unterscheidung zwischen verschiedenen Arten von Schlussfolgerung eingeführt werden.

2.1.1 Exkurs: Verschiedene Arten von Schlussfolgerungen

Es lassen sich drei Arten von Schlussfolgerungen unterscheiden:



Konzept 2-1: Arten von Schlussfolgerungen

1. Deduktive Schlussfolgerungen (Deduktion):

Hierbei handelt es sich um Schlussfolgerungen, die strikt *wahrheitserhaltend* sind, d.h. immer wenn die *Prämissen* (= die Aussagen, welche die Schlussfolgerung als gegeben annimmt) wahr sind, so ist auch die *Konklusion* (Der Satz, auf den geschlossen wird) wahr.

Umgekehrt muss bei Vorliegen einer deduktiven Schlussfolgerung im Falle einer falschen Konklusion mindestens eine der Prämissen falsch sein.

Die Korrektheit deduktiver Schlussfolgerungen ergibt sich rein durch die Form der Schlussfolgerung und nicht aufgrund des Inhalts der Aussagen.

Die Konklusion einer deduktiven Schlussfolgerung kann zu keine neuen Inhalte führen, welche nicht bereits implizit in den Prämissen vorhanden sind.

2. Induktive Schlussfolgerungen (Induktion):

Im Gegensatz zur Deduktion dienen induktive Schlüsse dazu, aufgrund von Gegebenem Neues zu erschliessen.

Der Prototyp induktiven Schlussfolgerns besteht in der Erschliessung eines gesetzmässigen Zusammenhanges aufgrund von vorliegenden Einzelbeobachtungen.

3. Abduktive Schlussfolgerungen (Abduktion):

Diese Art des Schliessens kann als ein Spezialfall des induktiven Schlussfolgerns betrachtet werden. Es wird auch als *diagnostisches Schliessen* bezeichnet.

Das Ziel des abduktiven Schlussfolgerns besteht darin, aus einer Menge vorhandener Erklärungen für einen vorliegenden Sachverhalt, diejenige zu finden, welche den Sachverhalt am besten erklärt.

Den klassische Fall von abduktiven Schlüssen bildet die Aufklärung eines Verbrechens mit mehreren Verdächtigen und natürlich ist es Sherlock Holmes, der durch seine genauen Beobachtungen und den daraus gezogenen abduktiven Schlussfolgerungen (und nicht deduktiven, wie er meinte) den Täter fängt.

Hier einige Beispiele zur Illustration der drei verschiedenen Arten des Schlussfolgerns:



Bsp. 2-1: Korrekte Deduktive Schlüsse:

Ein klassisches Beispiel für eine korrekte deduktive Schlussfolgerung ist der so genannte *Modus ponens*:

$$A \rightarrow B, A \therefore B$$

In Worten: Immer wenn *A* gilt, so gilt auch *B*. *A* ist wahr. Daher muss *B* wahr sein.

A → *B* und *A* sind die Prämissen und *B* ist die Konklusion.

Bsp.:

A: Die Schweiz befindet sich in Mitteleuropa.

B: Die Schweiz befindet sich in Europa.

A → *B:* Wenn die Schweiz sich in Mitteleuropa befindet, dann befindet die Schweiz sich in Europa.



Bemerkung zur Notation:

Die Symbole *A* und *B* sind so genannte Aussagenvariablen (Da es sich um Variablen handelt, werden sie im Text kursiv gesetzt). Sie repräsentieren irgendwelche Aussagen, die entweder wahr oder falsch sein können. Man kann daher – wie im obigen Beispiel – *A* und *B* durch konkrete Sätze ersetzen. Das Ergebnis eines Ausdrucks wie *A* → *B* ist wiederum ein Satz, der entweder wahr oder falsch ist.

Das Symbol \rightarrow ist ein so genanntes *logisches Konnektiv*, auch *Junktor* genannt (Deutsch: *logische Verknüpfung*). Konkret handelt es sich um die *materielle Implikation*. Diese wird umgangssprachlich mit »Wenn ... dann« wiedergegeben.

Diese Art der sprachlichen Wiedergabe verdeckt einen wichtigen Aspekt, nämlich, dass die Implikation $A \rightarrow B$ immer wahr ist, wenn A falsch ist (gleichgültig ob B wahr oder falsch ist). Demgemäß ist die Aussage »Wenn es auf Mars Menschen gibt, ist die Erde eine Scheibe« wahr (unter der Annahme, dass es auf dem Mars keine Menschen gibt).

Ist hingegen A wahr, so muss auch B wahr sein, wenn die Aussage $A \rightarrow B$ wahr sein soll.

Das Beispiel zeigt auch, dass die beiden Aussagen A und B keinen inhaltlichen Zusammenhang aufweisen müssen. Der Wahrheitswert von $A \rightarrow B$ ergibt sich nur aus den Wahrheitswerten der involvierten Aussagen: $A \rightarrow B$ ist nur dann falsch, wenn A wahr und B falsch ist, in allen anderen Fällen ist $A \rightarrow B$ wahr.

Das Symbol \therefore repräsentiert die Anweisung, dass von den *Prämissen* (= die Menge der Aussagen vor dem Symbol \therefore) auf die *Konklusion* (die Aussage nach dem Symbol \therefore) geschlossen werden darf.

Ein zweites Beispiel für eine korrekte deduktive Schlussfolgerung ist der so genannte *Modus tollens*:

$$A \rightarrow B, \neg B \therefore \neg A$$

In Worten: Immer wenn A gilt so gilt auch B . B gilt nicht. Folglich kann auch A nicht wahr sein.

Bsp.:

A : Die Schweiz befindet sich in Mitteleuropa.

$\neg A$: Die Schweiz befindet sich nicht in Mitteleuropa.

B : Die Schweiz befindet sich in Europa.

$\neg B$: Die Schweiz befindet sich nicht in Europa.

$A \rightarrow B$: Wenn die Schweiz sich in Mitteleuropa befindet, dann befindet die Schweiz sich in Europa.



Bemerkung zur Notation:

Das Symbol \neg repräsentiert die Negation: Immer wenn eine Aussage A wahr ist, so ist $\neg A$ falsch und umgekehrt.



Bsp. 2-2: Fehlerhafte deduktive Schlüsse:

Zwei typische Beispiele für fehlerhafte Schlüsse sind:

1. *Die Verneinung des Antezedens:*

$$A \rightarrow B, \neg A \therefore \neg B$$

2. Die Bejahung der Konsequenz:

$$A \rightarrow B, B \therefore A$$

Die folgende Unterscheidung ist ein zentraler Bestandteil der wissenschaftlichen Sprechweise und sollte daher verstanden werden.



Konzept 2-2: Notwendige und hinreichende Bedingung

Gegeben: Der Satz $A \rightarrow B$ (wobei A und B konkrete Sätze repräsentieren).

Die durch A spezifizierte Bedingung ist eine *hinreichende Bedingung* für die durch den Satz B spezifizierte Bedingung (kurz: A ist eine *hinreichende Bedingung* für B).

Die durch B spezifizierte Bedingung ist eine *notwendige Bedingung* für die durch den Satz A spezifizierte Bedingung (kurz: B ist eine *notwendige Bedingung* für A).

A ist *hinreichend* für B , da bei Vorliegen von A (d.h. A ist wahr) auch B vorliegen muss (gemäß modus ponens).

B ist *notwendig* für A , da bei Abwesenheit von B auch A abwesend sein muss (gemäß modus tollens).



Bsp. 2-3: Notwendige und hinreichende Bedingung:

Betrachte den Satz: Falls die StudentIn S einen Master in Psychologie hat, so hat sie auch einen Bachelor in Psychologie.

Die Aussage lässt sich als Konditionalsatz: $A \rightarrow B$ interpretieren, mit den Teilsätzen:

A : S hat einen Master in Psychologie.

B : S hat einen Bachelor in Psychologie.

Somit gilt:

Der Master in Psychologie ist eine *hinreichende Bedingung* für den Bachelor in Psychologie, während der letzterer eine *notwendige Bedingung* für den Master in Psychologie darstellt.



Bemerkung zur Zielsetzung der Logik:

Es mag befremdlich erscheinen, dass die Implikation $A \rightarrow B$ immer wahr ist, wenn die Aussage A falsch ist. Der Grund hierfür lautet wie folgt:

Die Logik beschäftigt sich nicht mit der Wahrheit von Aussagen, sondern mit der *formalen Gültigkeit* von *Schlussfolgerungen*. Im Zentrum steht daher die Frage, ob die Beziehung zwischen den Prämissen und der Konklusion derart beschaffen ist, dass die Konklusion notwendig aus den Prämissen folgt.

Dies bedeutet, dass niemals aus wahren Prämissen eine falsche Konklusion folgen kann. Derartige Schlussfolgerungen müssen strikt ausgeschlossen werden. Umgekehrt kann aus falschen Prämissen eine wahre oder auch eine falsche Aussage folgen.

Bevor wir das Gebiet der deduktiven Schlüsse endgültig verlassen und uns den beiden anderen Arten von Schlussfolgerungen zuwenden, noch eine kurze Bemerkung zur Wichtigkeit deduktiver Schlussfolgerungen.



Bemerkung zur Bedeutung deduktiver Schlussfolgerungen:

Auf den ersten Blick scheinen deduktive Schlussfolgerungen ziemlich nutzlos zu sein, da – wie oben in Konzept 2-1 ausgeführt – die Konklusion nichts enthält, was nicht schon implizit in den Prämissen enthalten ist.

Dieser Eindruck täuscht, denn in vielen Fällen ist die Anzahl der Prämissen hoch und es ist unklar, welche Konsequenzen aus diesen folgen. Eine formale Herleitung kann zeigen, dass eine Konklusion tatsächlich aus den Prämissen folgt.

Derartige Probleme finden sich z.B. in der Gesetzgebung, wo es immer wieder vorkommt, dass ein neues Gesetz im Verbund mit anderen unerwünschte Konsequenzen hat, welche aber nicht sofort bemerkt werden (Ein Beispiel betrifft die Schaffung unerwünschter »Steuerschlupflöcher« in der Steuergesetzgebung).

Wir wenden uns nun den induktiven Schlussfolgerungen zu.



Bsp. 2-4: Induktive Schlüsse:

Hier eine Liste von Beispielen der wichtigsten Formen von induktiven Schlussfolgerungen.

1. Schlussfolgerungen von Einzelbeobachtungen auf eine Gesetzmässigkeit:

Alle von mir beobachteten Schwäne waren weiss.

∴ Alle Schwäne sind weiss.

2. Statistische Verallgemeinerungen:

In einer an der Universität Fribourg erhobenen Stichproben ($N=100$) waren 90% der Studentinnen unverheiratet.

∴ 90% der Fribourger Studentinnen sind unverheiratet.

3. Analogieschlüsse:

Ein Experiment mit Ratten zeigte, dass die Substanz X456 ab einer Menge von 100 mg Krebs auslöste.

∴ Die Substanz löst auch beim Menschen Krebs aus.

4. Kausalschlüsse:

Wenn Personen mit Kreditkarte einkaufen, so achten sie weniger auf die Kosten.

∴ Kreditkarten verführen zum Geldausgeben.

Hier zwei Beispiele für abduktive Schlussfolgerungen:



Bsp. 2-5: Abduktive Schlussfolgerungen I:

Die Personen, welche den Film gesehen haben, waren nachher »geladen«.

Mögliche Erklärungen (Ursachen):

1. Die im Film dargestellte Gewalt erhöhte die Aggression der Zuschauer.
2. Die im Film gezeigte Ausbeutung einzelner durch staatliche Organe machte die Leute aggressiv.
3. Die schamlose Bereicherung einzelner der dargestellten Personen führte zu einer erhöhten Aggression bei den Zuschauern.

Abduktive Schlussfolgerung:

Die letzte Erklärung ist die plausibelste.



Bsp. 2-6: Abduktive Schlussfolgerungen II:

Hans gab Max auf der Party einen »Stoss«.

Mögliche Erklärungen (Ursachen):

1. *Person*: Hans ist ein Mensch der leicht zu Gewalttätigkeiten neigt.
2. *Objekt*: Max hatte Hans provoziert.
3. *Situation*: Auf der Party waren zu viele Leute auf zu engem Platz. Dieses »Crowding« führte zur erhöhten Aggressivität bei den Gästen.

Abduktive Schlussfolgerung:

Die letzte Erklärung ist die angemessenste.

Bemerkung:

Auf den ersten Blick scheint es zwischen Kausalschlüssen (vgl. Bsp. 2-4) und abduktiven Schlussfolgerungen keinen klar erkennbaren Unterschied zu geben. Dem ist jedoch nicht so.

- Im Falle von kausalen Schlüssen geht es um die *Erschließung eines kausaler Zusammenhanges* zwischen möglicher Ursache (Verwendung einer Kreditkarte) und Wirkung (vermehrtes Geldausgeben). Im Zentrum des Interesses steht hier die Problematik, ob es überhaupt einen kausalen Zusammenhang gibt und falls ja, über welchen kausalen Mechanismus die Ursache zu einer Wirkung führt.

- Im Falle von abduktiven Schlüssen, liegt eine völlig andere Situation vor. Hier geht es nicht um die Erschliessung neuer Ursachen bzw. kausalen Mechanismen. Im Zentrum steht vielmehr die Frage, welche der möglichen Ursachen, die bereits als mögliche Kandidaten identifiziert wurden, die aktuellen Ereignisse am besten erklärt.

Nach diesem kleinen Exkurs über verschiedene Arten von Schlussfolgerungen kehren wir wieder zu unserem Thema zurück.

2.1.2 Prinzipien der empiristisch-induktiven Charakterisierung der Wissenschaften

Der empiristischen Konzeption liegen zwei eng verwandte Prinzipien zugrunde:



Prinzip 2-1: Zentrales Prinzip des Empirismus:

Jede Art von Erkenntnis über die Welt ergibt sich aus der Erfahrung. Im Speziellen ergeben sich wissenschaftliche Gesetze und Theorien durch Anwendung der Induktion. Hierbei werden aufgrund von Einzelbeobachtungen via induktive Schlüsse Gesetze und Theorien erzeugt.

Gemäss Prinzip 2-1 wählen korrekt operierende Wissenschaftler folgende Vorgehensweise. Sie versuchen, die Naturvorgänge möglichst genau zu beobachten und induzieren aufgrund von beobachteten Regelmässigkeiten wissenschaftliche Gesetze und Theorien.



Bsp.2-7: Anwendung der Induktion I: Klassische Konditionierung:

Pawlow bemerkte, dass der von ihm beobachtete Hund Speichel absonderte sobald eine Glocke ertönte, die zuvor über längere Zeit hinweg das Fressen für den Hund angekündigt hatte.

Er schloss daher auf folgenden gesetzesmässigen Zusammenhang:

Wird ein Reiz *CS* (Glocke) oft genug mit einem anderen Reiz *US*, der eine bestimmte Reaktion *CR* auslöst, gepaart, so führt das Vorhandensein des *CS* alleine (ohne die Präsentation von *US*) bereits zur *CR*.



Bemerkung zur Notation:

CS = konditionierter Stimulus;

US = unkonditionierter Stimulus;

CR = konditionierte Reaktion.



Bsp.2-8: *Anwendung der Induktion II: Zentrum für Buchstabenerkennung im Gehirn:*

Aufgrund von fMRI – Studien wurde festgestellt, dass beim Lesen eine bestimmte Region im linken unteren Schläfenlappen aktiviert wird (Dehaene, 2010).

Hieraus wird auf folgenden Zusammenhang geschlossen:
Die betreffende Region repräsentiert Buchstaben.



Bsp.2-9: *Anwendung der Induktion III: Untersuchungen zur Entdeckung von Zusammenhängen*

In klinischen und angewandten Bereichen der Psychologie werden oft empirische Untersuchungen durchgeführt, in denen möglichst viele Variablen gemessen werden. Das Ziel besteht darin, mögliche Zusammenhänge zwischen den Variablen aufzufinden, z.B. welche Größen einen Zusammenhang mit psychischem Wohlbefinden oder psychischer Gesundheit aufweisen.

Derartige Untersuchungen werden oft (etwas abfällig) als »Schrotschussuntersuchungen« bezeichnet, da – ähnlich wie beim Schiessen mit Schrot – durch die Einbeziehung einer grossen Anzahl von Variablen eine möglichst breite Streuung erzielt werden soll.

Diese Beispiele machen deutlich dass die Forschungstätigkeit in der Psychologie – zumindest teilweise – unter Heranziehung von Prinzip 2-1 verstanden werden kann.

Es sei nun noch auf ein zweites Prinzip verwiesen, welches charakteristisch für den späten Empirismus ist und als Teil der so genannten »received view« betrachtet wurde (Eine Zusammenfassung der Prinzipien des späten Empirismus findet man in Suppe, 1977, Seite 50-52).



Prinzip 2-2: *Trennung von theoretischer Sprache und Beobachtungssprache:*

Die folgenden beiden Sprachen sind strikt zu trennen:

- (a) Die *theoretische Sprache* enthält theoretische Terme, welche sich nicht auf direkt beobachtbare Entitäten beziehen.
- (b) Die *Beobachtungssprache* enthält nur Begriffe, welche sich auf direkt Beobachtbares beziehen.

Theoretische Begriffe erhalten ihre Bedeutung dadurch, dass sie sich entweder direkt auf Beobachtungsbegriffe reduzieren lassen, oder indirekt auf andere theoretische Konstrukte, die sich ihrerseits auf beobachtbare Größen zurückführen lassen.

Theoretische Begriffe, für welche dies nicht gilt (wie z.B. *Triebenergie* oder *energetische Körperzonen*) sind bedeutungslos.

Das folgende Beispiel verdeutlicht die Unterscheidung zwischen theoretischen Konstrukten und Beobachtungen, sowie deren Beziehung:



Bsp.2-10: Facetten des Konzepts der Kontrollüberzeugung
(nach Krampen, 1991)

Gegeben: Das theoretische Konstrukt der (subjektiven) *Kontrollüberzeugung*, bestehend aus den 4 theoretischen Teilkonstrukten:

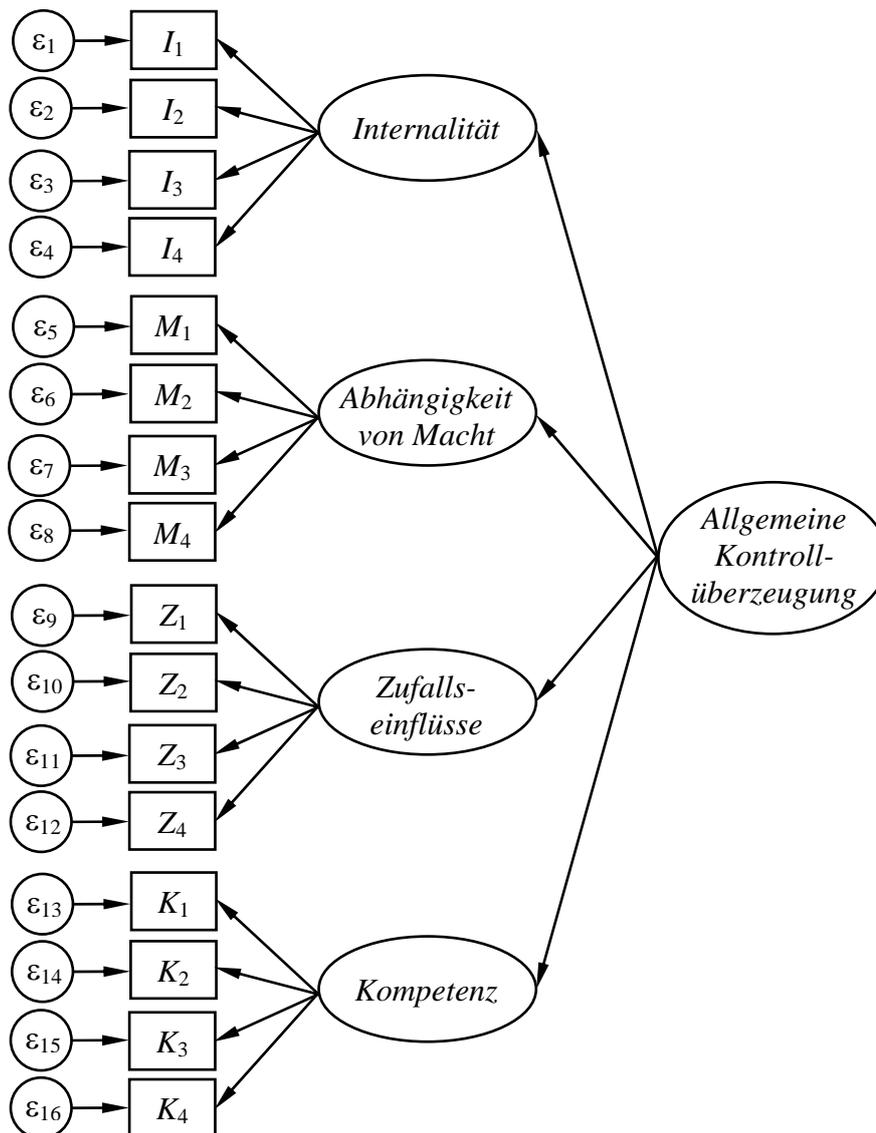


Abb. 2-1: Modell der Kontrollüberzeugung.

1. *Internalität-Externalität*: Überzeugung bezüglich des Locus von (Miss-) Erfolg.
2. *Abhängigkeit von Macht*: Einschätzung der Beeinflussbarkeit durch äussere Macht.
3. *Zufallseinflüsse*: Einschätzung der Bedeutung des Zufalls.

4. *Kompetenz*: Einschätzung der eigenen Kompetenz.

In Abb. 2-1 sind die theoretischen Konstrukte durch Ellipsen symbolisiert.

Jede der 4 Teilkonstrukte wird auf 4 Subskalen eines Fragebogens gemessen. Diese Messungen bilden die konkreten Beobachtungen. Sie sind in Abb. 2-1 durch die Rechtecke symbolisiert.

Die kleinen mit ε_i bezeichneten Kreise symbolisieren so genannte *Fehlerterme*. Sie repräsentieren Einflüsse auf die Messungen, welche durch die Konstrukte im Modell nicht erklärt werden können.

Die Teilkonstrukte der *Kontrollüberzeugung* stehen in direkter Beziehung zu den beobachteten Grössen. Dies wird in Abb. 2-1 durch die Pfeile repräsentiert. Diese drücken eine kausale Beziehung aus, d. h. jedes Teilkonzept hat einen direkten Einfluss auf die Antwort der Personen auf die zugehörigen Fragen, welche zur Messung des Teilkonzeptes gedacht sind (nicht aber auf die anderen). Sie haben daher einen messbaren Effekt und erhalten so ihre Bedeutung.

Das Konstrukt der *Allgemeinen Kontrollüberzeugung* besitzt hingegen keine direkte Beziehung zu den konkreten Messungen. Da dieses Konstrukt jedoch direkt mit den Teilkonstrukten verbunden ist (indem es diese kausal beeinflusst), erhält es durch diese ihre Bedeutung.

Das in Abb. 2-1 dargestellte Modell ist ein so genanntes faktorenanalytisches Modell (wobei im aktuellen Falle die Beziehungen zwischen den verschiedenen Grössen kausal interpretiert wurden).

Die Methode der (explorativen) Faktorenanalyse findet in der Psychologie eine breite Anwendung. Sie passt einerseits sehr gut zur induktiv-empiristischen Konzeption, kann aber auch als eine der am meisten missbrauchten statistischen Methoden betrachtet werden. Der folgende kleine Exkurs beleuchtet diese Aspekte näher.

2.1.3 Exkurs: Zu Bedeutung und Missbrauch der Faktorenanalyse in der psychologischen Forschung

Die Faktorenanalyse existiert in zwei grundlegenden Spielarten: Als *explorative* oder als *konfirmative* Technik. Im Folgenden betrachten wir nur die explorative Variante (zur konfirmativen Form, siehe Abschnitt xxx). Wenn daher in der weiteren Darstellung von »Faktorenanalyse« die Rede ist, so ist immer nur die explorative Version gemeint.

Die explorative Faktorenanalyse (EFA) wird im Allgemeinen für die Entwicklung von psychologischen Tests – wie Intelligenz- oder Persönlichkeitstests – verwendet.



Method 2-1: Explorative Faktorenanalyse (EFA):

Die Anwendung der explorativen Faktorenanalyse (EFA) vollzieht sich in drei Schritten:

1. *Datensammlung:*

Es werden möglichst viele Variablen gemessen, welche für den aktuellen Bereich relevant sein könnten (Man beachte hier den Zusammenhang mit den in Bsp.2-9 besprochenen Schrotschussuntersuchungen). Zwischen allen Variablen werden die Kovarianzen oder Korrelationen berechnet und zu einer Kovarianz- bzw. Korrelationsmatrix zusammengefasst.

2. *Faktorenextraktion:*

Aufgrund der Kovarianz- oder Korrelationsmatrix der Variablen werden mit Hilfe eines statistischen Verfahrens (im Falle der so genannten Maximum-likelihood EFA) oder mittels eines »pseudostatistischen« (wie z.B. der sehr häufig verwendeten Hauptfaktorenanalyse) eine möglichst geringe Anzahl Faktoren extrahiert.

3. *Finden einer einfachen Faktorenstruktur:*

Mit Hilfe einer so genannten Faktorenrotation wird eine einfache Faktorenstruktur gesucht. Dies bedeutet, dass die erhobenen Grössen jeweils nur wenige Faktoren (im Idealfall auf einen) laden (vgl. das Modell in Abb. 2-1, wo jede beobachtete Grösse nur von einem einzelnen Faktor beeinflusst wird). Variablen, die auf keinen Faktor laden werden eliminiert.



Bemerkung zur Hauptfaktorenanalyse:

Oben wurde die Hauptfaktorenanalyse, welche in der Psychologie die am meisten verwendete faktorenanalytische Methode darstellt als »pseudostatistisches« Verfahren bezeichnet. Der Grund für diese Wahl der Terminologie findet sich in der Tatsache, dass die Methode keinerlei Verteilungsannahmen macht. Daher gibt es auch keinen statistischen Test zur Prüfung, wie gut das Modell die Daten erklärt.

Im Falle der Maximum-likelihood EFA gibt es einen statistischen Test, der Auskunft über die Modellgüte gibt.

Die EFA kann als eine Methode betrachtet werden, welche aufgrund von konkreten Beobachtungen theoretische Konstrukte induziert. Sie passt daher in idealer Weise zur induktiv-empiristischen Konzeption der Wissenschaft.

An diesem Punkt setzt aber auch die Kritik bezüglich des Missbrauchs der EFA an, die sich wie folgt zusammenfassen lässt:

Die EFA ersetzt das Nachdenken über theoretische (mentale) Konstrukte und deren Zusammenhänge durch eine mechanische Prozedur. Dies führt zu theoretisch uninteressanten und unergiebigen Theorien.

Das folgende Beispiel dient zur Illustration dieser Kritik.



Bsp.2-11: Faktorenanalyse und Persönlichkeitsforschung:

In dem folgenden Beispiel karikiert Seymour Epstein (1994) die Persönlichkeitsforschung. Speziell geht es ihm um die Forschung zu den so genannten Big Five (siehe z.B. McCrae & John, 1990).

Das Hauptergebnis dieser Persönlichkeitstheorie besagt, dass die Persönlichkeit aus 5 Faktoren aufgebaut ist, wobei zu jedem Faktor wiederum zahlreiche Unterfaktoren (Facetten) vorliegen. Die 5 Persönlichkeitsfaktoren sind:

1. Extraversion
2. Verträglichkeit
3. Gewissenhaftigkeit
4. Neurotizismus
5. Offenheit für neue Erfahrungen

Hier nun Epsteins Geschichte:

Es war einmal ein Psychologe mit Namen Sam. Der entschied sich, Autos anstelle von Leuten zu untersuchen. Er glaubte es sei wissenschaftlicher, Autos zu studieren, da diese besser kontrollierbar und einfacher zu verstehen seien als Menschen. Sam wies darauf hin, dass Autos viel mit Menschen gemeinsam hätten, wie Bewegung, Essen (d.h. Aufnahme von Benzin), Atmung (d.h. Aufnahme von Sauerstoff) Ausscheidung (d.h. Produktion von Abgasen), individuelle Persönlichkeiten und Fehlfunktionen. Daher sollte das Wissen, welches er aus dem Studium der Autos erlangte, auf Leute transferierbar sein. »Eins nach dem anderen«, sagte er sich, »sobald ich gelernt habe, die Persönlichkeiten von Autos exakt darzustellen, kann ich mich an die Menschen heranwagen.«

Sam begann nun die fundamentalen Merkmale von Autos zu ermitteln. Mit Hilfe einer Faktorenanalyse enthüllte er fünf unabhängige Merkmale: Farbe (schwarz, weiss, rot, blau, etc.), Typ (Kombi, Sportcoupé, Limousine, Freizeitwagen, etc.), Grösse (kompakt, mittel, volle Grösse), maximale Geschwindigkeit und Robustheit (Servicestatistik).

Die Ergebnisse waren höchst eindrucksvoll. Sachverständige konnten die Klassifikation reliabel verwenden und diese war umfassend (indem sie alle Autotypen einschloss). Das System galt innerhalb einer Kultur und auch zwischen verschiedenen Kulturen und alle Arten von interessanten Beobachtungen korrelierten damit. Zum Beispiel bevorzugten grosse Familien mit Haustieren Kombis, während Singles mit einem grossem Bedürfnis nach Stimulation eher Sportwagen präferieren.

Die Lösung war allerdings nicht so einfach wie sie scheint. Sam musste hart arbeiten, um schlussendlich die Grundeigenschaften zu ermitteln, aber Dank hochleistungsfähiger Rechner war er schliesslich in der Lage, die Aufgabe zu vollenden. Hier ein Beispiel für die sich ergebenden Schwierigkeiten: Sam hatte zuvor andere Dimensionen einbezogen, die er für wichtig hielt. Die Faktorenanalyse belehrte ihn jedoch, dass diese als Grunddimensionen wertlos – weil nicht unabhängig – seien. Dazu gehör-

ten wissenschaftlich wertlose Variablen wie Gewicht, Länge, Breite und Höhe. Nach der faktorenanalytischen Behandlung ergab sich hieraus ein einzelner Faktor: Dichte. Zuerst erschien dies als elegante Lösung, aber dann fand Sam, dass Dichte nicht von der Grösse unabhängig war und letztere erklärte mehr Varianz als Dichte (in einer Regressionsgleichung). Somit hatte er keine andere Wahl als die Variable Dichte fallen zu lassen. Sam war sehr stolz auf den Grad an Organisation, den er in die Klassifikation von Autos mit nur fünf Faktoren hatte bringen können. Eines Tages als er sich gerade auf dem Weg befand, um seine neuesten Ergebnisse an einem psychologischen Kongress vorzustellen, hatte er eine Autopanne. Unglücklicherweise hatte er nicht die geringste Idee, was zu tun sei. Er wusste nichts darüber, wie ein Auto funktionierte. In der Tat hatte er nie unter die Haube geschaut. Was sich darunter befand, hatte er immer mit dem tiefen dunklen Unbewussten des Menschen gleichgesetzt und hielt es deshalb für unwichtig. Er musste sich also mit dem begnügen, was er hatte, d.h. mit allen Arten von Forschungsergebnissen darüber, wie die grundlegenden Eigenschaften eines Autos mit dem Auftreten einer Panne korreliert sind. Farbe weist gewöhnlich keine signifikanten Beziehungen mit der Pannenstatistik auf, obwohl es eine nichtsignifikante Tendenz gibt, dass Leute mit roten Autos mehr Unfälle produzieren. Robustheit war hingegen eine relevante Grösse und sagte selbstverständlich zukünftige Pannen überzufällig vorher. Nun verstand er warum sein Auto eine Panne hatte: Es hatte eine geringe Ausprägung auf dem Faktor Robustheit. Es kam ihm der Gedanke, dass dies leicht auf menschliches Verhalten übertragen werden könnte, indem man Neurosen an die Stelle von Pannen setzte. Nun konnte er endlich erklären, warum Menschen psychologische Symptome hatten: Weil sie neurotisch waren. Leute entwickeln Symptome, denn sie haben eine hohe Ausprägung auf der Dimension Neurotizismus, ebenso wie Autos Pannen haben, weil sie einen geringen Wert auf dem Merkmal Robustheit besitzen. Unglücklicherweise teilte sein Auto Sams Enthusiasmus über diese neue Erkenntnis nicht und rührte sich auch weiterhin keinen Zentimeter.

Letztendlich entschied Sam, dass Autos doch zu kompliziert seien. Er wandte sich daher wieder dem Studium der Menschen zu. Er hielt den Traum aufrecht, dass er – durch Extraktion der einzelnen Faktoren mittels Faktorenanalyse und Aufdeckung weiterer Facetten – schliesslich in der Lage sein werde, die Kennzeichnung jeder Art von menschlichem Befinden zu ermitteln. Dies bedeutete für ihn vollständiges Verstehen.

(Nach Epstein, 1994, Seite 120-121).

Die Kritik Epsteins läuft darauf hinaus, dass uns die Persönlichkeitsforschung im Bereich der Big-Five keine grundlegenden Erkenntnisse über die Funktion der menschlichen Persönlichkeit liefert. Die Ergebnisse können höchstens als eine Prototheorie betrachtet werden.

Bemerkung:



Die Tatsache, dass die Big-Five Persönlichkeitsforschung keine grundlegenden Erkenntnisse zur menschlichen Persönlichkeit liefert, bedeutet nicht, dass sie keinen praktischen Nutzen hat.

Der Nutzen liegt darin, dass die Persönlichkeitsfaktoren einen prognostischen Wert bezüglich wichtiger Lebensbereiche – wie Lebenszufriedenheit oder Berufserfolg – haben können (Siehe die Geschichte von Sam)

Nach dieser Darstellung der wichtigsten Prinzipien des Empirismus und deren Anwendungen im Bereich der psychologischen Forschung wenden wir uns der zweiten Konzeption der Wissenschaft zu, die auch heute noch einen grossen Einfluss aufweist.

2.2 Die deduktiv-falsifikationistische Konzeption der Wissenschaft

Für Karl Popper (1902–1994) sind die beiden Prinzipien der induktiv-empiristisch Konzeption der Wissenschaft fundamental falsch. Dies betrifft sowohl deren Rolle als mögliche *normative Prinzipien*, d.h. als Vorschriften, wie Wissenschaft ablaufen soll, als auch deren Bedeutung als *deskriptive Prinzipien*, d.h. als Beschreibung der tatsächlichen wissenschaftlichen Praxis.

2.2.1 Poppers Kritik an der induktiv-empiristischen Konzeption

Betrachten wir nun Poppers Kritik an den Prinzipien der induktiv-empiristischen Auffassung im Detail.

2.2.1.1 POPPERS KRITIK DER INDUKTION

Den Ausgangspunkt von Poppers Überlegungen bildet das Induktionsproblem.



Konzept 2-3: Das Induktionsproblem:

Das Induktionsproblem lässt sich durch die folgenden beiden Frage charakterisieren:

1. Wie lassen sich induktive Schlussfolgerungen rechtfertigen?
2. Können induktive Schlussfolgerungen ähnlich wie deduktive Schlüsse gerechtfertigt werden?

Poppers Antwort auf diese Frage ist eindeutig: Induktive Schlussfolgerungen können nicht gerechtfertigt werden. Aus diesem Grunde macht es wenig Sinn, Induktion als normatives Prinzip der Erkenntnisgewinnung zu betrachten. Die Wissenschaft sollte daher vollständig auf diese Methode verzichten.



Bemerkung:

Der Skeptizismus bezüglich der Rechtfertigung jeglicher Art von induktiver Schlussfolgerung wurde bereits von dem schottischen Philosophen David Hume (1711–1776) eindringlich formuliert. Im Gegensatz zu Popper führte dies bei Hume jedoch nicht zur Aufgabe der Induktion als nützliches Prinzip.

Für Popper ist die Induktion nicht nur als normatives Prinzip unge-rechtfertigt, es beschreibt auch in keiner Weise, wie sich die Theoriebildung in den Wissenschaft vollzieht. Für ihn beginnt der Prozess der Theoriebildung mit einem *Problem*, mit dem sich eine Forscherin konfrontiert sieht. Die Bildung der Theorie wird als Prozess der Problemlösung betrachtet. In vielen Fällen entsteht das Problem aufgrund von konkreten – meist überraschenden – Beobachtungen, dies ist jedoch nicht immer der Fall. Die eigentliche empirische Forschung beginnt jedoch erst, nachdem das Problem aufgetaucht ist und verschiedene mögliche Theorien als Lösungsversuche entwickelt wurden.

Betrachten wir unter diesem Gesichtspunkt nochmals die oben gegebenen Beispiele:



Bsp.2-12: *Klassische Konditionierung (Fortsetzung von Bsp.2-7)*

Pawlow bemerkte, dass der von ihm beobachtete Hund Speichel absonderte, sobald eine Glocke ertönte, die zuvor über längere Zeit hinweg das Fressen für den Hund angekündigt hatte.

Dies war für ihn eine überraschende Beobachtung und er stellte sich die Frage, welcher Mechanismus dahinter steckt. Er hypostasierte, dass der Hund eine Assoziation zwischen CS (Glocke) und US (Futter) gebildet hatte.



Bsp.2-13: *Zentrum für Buchstabenerkennung im Gehirn (Fortsetzung von Bsp.2-8):*

Aufgrund von fMRI – Studien wurde festgestellt, dass beim Lesen eine bestimmte Region im linken unteren Schläfenlappen aktiviert wird (Dehaene, 2010).

Hieraus ergeben sich eine Reihe von Fragen, z.B.:

- Wie werden die Buchstaben in dieser Region repräsentiert (invariant bezüglich Rotation, Grösse und Verschiebung?)
- Welche Funktion hat diese Region im Gesamtkontext des Lesens?



Bsp.2-14: *Untersuchungen zur Entdeckung von Zusammenhängen (Fortsetzung von Bsp.2-9):*

Findet man in einer explorativen Studie Zusammenhänge zwischen verschiedenen Variablen, so stellt sich sofort die Frage nach den kausalen Wirkmechanismen, welche für den beobachteten Zusammenhang verantwortlich sind.

Nehmen wir an, unsere Forscherin findet einen Zusammenhang zwischen partnerschaftsexternem Stress (z.B. Stress am Arbeitsplatz) und der erlebten Beziehungsqualität bei Paaren.

Hier erhebt sich nun die Frage, ob dieser externe Stress direkt auf die Qualität der Partnerschaft wirkt oder über partnerschaftsinternen Stress vermittelt wird (siehe Abb. 2-2).

Eine weitere Frage betrifft die unterschiedliche Wirkung auf die beiden Partner. Hat der externe Stress des männlichen Partners eine stärkere Wirkung als jener der Partnerin?

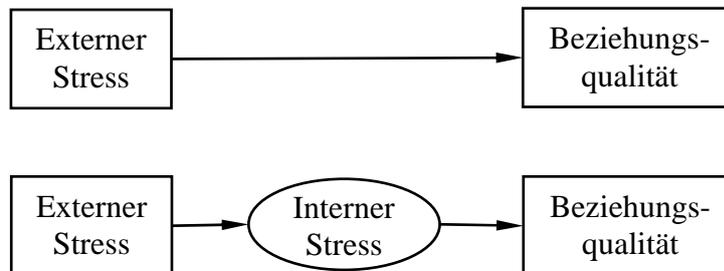


Abb. 2-2: Kausale Modelle zur Wirkung von partnerschaftsexternem Stress auf die Beziehungsqualität.

Die wirkliche Forschung beginnt erst mit der *gezielten Untersuchung* dieser und ähnlicher Fragestellungen. Hierbei werden gezielt nur noch die theoretisch relevanten Grössen in die Untersuchung einbezogen und nicht mehr alle möglichen.

Für Popper vollzieht sich die Theoriebildung in der Regel auch nicht durch Verallgemeinerungen von Beobachtungen. Dies ist unmöglich, da fruchtbare Theorien theoretische Entitäten beinhalten, die nicht direkt beobachtbar sind.



Bsp.2-15: Theoretische Entitäten in der Alltagspsychologie und in der wissenschaftlichen Psychologie

Sowohl die Alltagspsychologie wie die wissenschaftliche Psychologie »wimmelt« nur so von theoretischen Entitäten, welche dazu verwendet werden beobachtetes Verhalten zu erklären, zu verstehen und vorherzusagen.

Als theoretische Entitäten fungieren in der Psychologie mentale Zustände, Dispositionen und Prozesse.

- Die Alltagspsychologie verwendet kurzfristige *mentale Zustände*, wie Emotionen (Wut, Angst, Freude, etc.) aber auch länger andauernde *mentale Dispositionen*, wie Eifersucht, Verliebtheit, Aggressivität, etc.
- Die wissenschaftliche Psychologie hat diese mentalen Zustände und Dispositionen ebenfalls zum Gegenstand. Sie versucht diese genauer zu erfassen. Hinzu kommen noch weitere mentale Konstrukte, welche sich auf weitere mentale Repräsentationen und Prozesse beziehen, z.B. Semantische Netzwerke, Prozesse der visuellen Verarbeitung von Buchstaben, oder Urteilsheuristiken.

Die Bildung und Verwerfung von Theorien in der Psychologie wird in Kapitel 3 näher beleuchtet. Wir wenden uns als nächstes einer weiteren Kritik Poppers an der induktiv-empiristischen Konzeption zu.

2.2.1.2 POPPERS KRITIK AN DER TRENNUNG VON BEOBACHTUNGS- UND THEORETISCHER SPRACHE

Für Popper gibt es keine strikte Trennung zwischen Beobachtungs- und theoretischer Sprache, wie folgendes Zitat eindringlich belegt:

»Unsere Sprache ist von Theorien durchsetzt: **es gibt keine reinen Beobachtungssätze**. [...] Sogar in einer sogenannten „phänomenalen“ Sprache, die etwa „jetzt hier rot“ zulässt, würde das Wort „jetzt“ eine (rudimentäre) Theorie der Zeit implizieren; das Wort „hier“ eine Theorie des Raumes; und das Wort „rot“ eine Theorie der Farben.

[...]

Es gibt keine reinen Beobachtungen: sie sind von Theorien durchsetzt und werden von Problemen und Theorien geleitet.« (Popper, 1989; S. 76 [Hervorhebungen im Original]).

Man beachte, dass für Popper eine Trennung der beiden Arten von Sprachen deshalb problematisch ist, weil auch unsere Beobachtungen theoriendurchtränkt sind und die Idee einer reinen Beobachtung somit hinfällig ist.

Betrachtet man diese Problematik vom Standpunkt einer Forscherin, so wird dies sofort offensichtlich. Denn schon am Beginn jeder Untersuchung hat unsere Psychologin bestimmte Ziele und Annahmen im Kopf. Dies ist auch dann der Falle wenn sie eine sehr breit angelegt Explorationsstudie durchführen will. Denn auch in diesem Fall muss Sie die zu messende Größen festlegen und diese Auswahl wird von ihren Vorannahmen über die Relevanz der verschiedenen Variablen mit bestimmt. Eine theorielose und völlig unvoreingenommene Forschung, wie es manchen Empiristen vorschwebte, ist daher grundsätzlich unmöglich.



Bemerkung:

Man beachte die Übereinstimmung zwischen Popper und Quine, der die physikalischen Dinge als theoretische Setzungen betrachtet, die es uns ermöglichen Ordnung in den Strom unserer Erfahrungen zu bringen (vgl. das obige Zitat in Kapitel 1).

Die Theoriegeladenheit bzw. Beobachtungsnähe von Begriffen ist also keine Alles-oder-Nichts Phänomen, sondern eine kontinuierliche Eigenschaft: Begriffe sind mehr oder weniger beobachtungsnah.

John Watkins (1924–1999), ein Schüler von Popper stellte folgende Kriterien zur Unterscheidung von mehr oder weniger beobachtungsnahen Begriffen auf.



Prinzip 2-3: Kriterien zur Unterscheidung zwischen beobachtungsnahen versus theoriegeladenen Begriffen (Watkins, 1984):

1. Ein Begriff ist umso beobachtungsnäher, je einfacher es ist zu entscheiden, ob ein konkretes Objekt unter diesen fällt.

2. Ein Begriff ist umso beobachtungsnäher, je weniger Instrumentarium benötigt wird, um zu entscheiden, ob er auf ein konkretes Objekt zutrifft.
3. Ein Begriff ist umso beobachtungsnäher, je weniger Rückgriff auf eine Theorie benötigt wird, um ihn zu verstehen.

Im Allgemeinen sind beobachtungsnähere Konstrukte weniger problematisch, in dem Sinne, dass sich leichter eine Einigung über deren Existenz erzielen lässt. Es ist meist einfacher über konkret beobachtbares Verhalten, wie z.B. eine aggressive Handlung, Einigkeit zu erzielen als über eine zugrunde liegende Disposition, wie z.B. die Aggressivität einer Person.

Insgesamt werden daher Theorien und Gesetze einfacher zu beurteilen sein, wenn nur relativ beobachtungsnah theoretische Konstrukte involviert sind, gegenüber Theorien mit hoch abstrakten theoretischen Konzepten. Dieser Aspekt ist sowohl für die Beurteilung von wissenschaftlichen wie von pseudowissenschaftlichen Theorien und Praktiken relevant.



Bemerkung zur Notation:

Trotz der Tatsache, dass Begriffe mehr oder weniger beobachtungsnah sind, werden wir im Folgenden – der Bequemlichkeit halber – manchmal von theoretischen Begriffen und beobachtbaren Dingen sprechen. Wir nehmen also eine (willkürlich) gezogene Trennlinie zur Kategorisierung eines Objektes als theoretisch bzw. beobachtbar an.

Wir haben nun die zentralen Kritikpunkte Popper an der induktiv-empiristischen Konzeption der Wissenschaft behandelt. Bevor wir uns Poppers eigener Konzeption zuwenden, seien noch drei wichtige Prinzipien, welche Poppers *kritischem Rationalismus* zugrunde liegen, besprochen.

2.2.2 Poppers kritischer Rationalismus

Poppers Denken ist von drei wichtigen Prinzipien geleitet, die in enger Verbindung stehen.



Prinzip 2-4: *Unmöglichkeit voraussetzungsloser Erkenntnis:*
Jede Theorie beruht auf Annahmen, welche keiner weiteren Rechtfertigung unterzogen werden.

Begründung:

Die Rechtfertigung von Annahmen beruht auf weiteren Annahmen, welche selbst wiederum der Rechtfertigung bedürfen. Dies würde zu einem *unendlichen Regress* führen, d.h. einem unendlichen Prozess der Rückführung von Annahmen auf andere grundlegendere Annahmen.

Popper hat die in Prinzip 2-4 dargestellte Problematik in folgendem berühmten Zitat wunderbar ausgedrückt:

»Die Wissenschaft baut nicht auf Felsengrund. Es ist eher Sumpfland, über dem sich die kühne Konstruktion ihrer Theorien erhebt; sie ist ein Pfeilerbau, dessen Pfeiler sich von oben her in den Sumpf senken – aber nicht bis zu einem natürlichen gegebenen Grund. Denn nicht deshalb hört man auf die Pfeiler tiefer hineinzutreiben, weil man auf eine feste Schicht gestoßen ist: wenn man hofft, dass sie das Gebäude tragen werden beschließt man, sich vorläufig mit der Festigkeit der Pfeiler zu begnügen.« (Popper, 1989; S. 75-76).

Klarerweise sollten die Annahmen, auf denen ein Theoriengebäude beruht, möglichst unproblematisch sein, d.h. von Mitgliedern der wissenschaftlichen Gemeinschaft akzeptiert werden. Eine grundlegende Annahme der Psychologie lautet z.B. wie folgt:



Bsp. 2-16: Eine grundlegende Annahme der Psychologie:

Menschliches Erleben und Verhalten unterliegt bestimmten Gesetzmässigkeiten.

Falls diese Annahme nicht zutrifft, so ist einer wissenschaftlichen Psychologie die Grundlage entzogen. Man beachte jedoch, dass diese Annahme auch als Grundlage für den alltäglichen Umgang der Menschen untereinander dient, denn auch im Alltag wird stillschweigend vorausgesetzt, dass menschliches Verhalten bestimmte Regelmässigkeiten aufweist. Wie wäre sonst ein geregelter Umgang untereinander überhaupt möglich?

Für Popper liegt die einzig sinnvolle Antwort auf die in Prinzip 2-4 formulierte Tatsache in folgendem Prinzip:



Prinzip 2-5: Prinzip der Offenheit und permanenten Kritik:

Alle Annahmen bezüglich Theorien, Randbedingungen, Methoden oder logischen Voraussetzungen müssen jederzeit der Kritik zugänglich und damit revidierbar sein.

Der Prozess der Revision von Theorien und Annahmen ist daher niemals abgeschlossen.

Die Überlegungen, auf welchen Prinzip 2-4 und Prinzip 2-5 beruhen, lassen sich wie folgt zusammenfassen:

Um Forschung zu betreiben, benötige ich bestimmte Annahmen, die ich als vorläufig unproblematisch voraussetze. Alle diese Annahmen »lege ich – fairerweise – auf den Tisch«, damit sie der öffentlichen Kritik zugänglich werden. Sollte sich herausstellen, dass meine Annahmen der Kritik nicht standhalten, so muss ich sie revidieren und die notwendigen Konsequenzen für meine Theorien ziehen.



Bsp.2-17: Kritik der Grundlagen der Psychologie:

Innerhalb der Psychologie gab es immer wieder grundlegende Annahmen betreffende Diskussionen. Hier einige Beispiele:

1. Diskussion zu Annahmen über die Form psychologischer Theorien:
Inwieweit sollen Konzepte des Alltagspsychologie – wie Ziele, Wünsche, etc. – Eingang in die Theorien der Psychologie finden? So lehnten die Behavioristen jede Art von mentalen Konzepten als unwissenschaftlich ab. Für die Neurophilosophen Patricia und Paul Churchland sind sie als wissenschaftliche Konzepte wenig brauchbar, da Vorhersagen auf Basis dieser Konzepte nur eingeschränkt möglich und dabei noch ungenau sind.
2. Diskussion über die Bedeutung der Introspektion:
Inwieweit können mentale Prozesse via Introspektion erfasst werden? Die Gestaltpsychologie lehnte Erkenntnisse über mentale Prozesse, welche auf Introspektion beruhen, ab. Später wurde Introspektion in Form des so genannten *Lauten Denkens* zur Erfassung von Denkprozessen durch die Arbeiten von Newell und Simon wieder salonfähig. Die Grenzen der Introspektion wurden von Nisbett und Wilson (1977) demonstriert.
3. Diskussion über die Bedeutung statistischer Signifikanztests:
Seit vielen Jahren gibt es eine Diskussion über Sinn und Unsinn statistischer Signifikanztests (Harlow, Mulaik, & Steiger, J. L., 1997).
4. Diskussion über den Wert der explorativen Faktorenanalyse (Vergleiche Kapitel 2.1.3).

Ein weiteres Prinzip, welches als Pendant zu Prinzip 2-4 betrachtet werden kann, lautet wie folgt:



Prinzip 2-6: Voraussetzung für einen rationalen Diskurs

Ein rationaler Diskurs zwischen Personen benötigt als Voraussetzung die Anerkennung minimaler Rationalitätsstandards durch die Teilnehmer.

Zu diesen Voraussetzungen gehören unter anderem:

- Anerkennung elementarer logischer Prinzipien,
- Anerkennung des Prinzips, dass nur das bessere Argument zählt, während Gewalt und Status irrelevant sind.

Wir kommen als nächstes zu Poppers Konzeption der Wissenschaft.

2.2.3 Poppers deduktiv-falsifikationistische Konzeption der Wissenschaft

Im Folgenden diskutieren wir die zentralen Prinzipien, welche Poppers Konzeption der Wissenschaft bestimmen.



Prinzip 2-7: Unterscheidung zwischen Entdeckungs- und Begründungszusammenhang (Reichenbach, 1938)

1. Der *Entdeckungszusammenhang* betrifft die Entstehung von neuen Theorien. Hier geht es um Fragen wie:
 - Wie entstehen neue Theorien und Modelle?
 - Was begünstigt die Entstehung guter Theorien?
 - Gibt es allgemeine Prinzipien, welche der Bildung neuer Theorien zugrunde liegen?

Für Popper gehört der gesamte Komplex der Entstehung von Theorien in das Gebiet der Psychologie und nicht in jenes der Wissenschaftsphilosophie.

2. Der *Begründungszusammenhang* betrifft die Rechtfertigung und Prüfung von Theorien. Hier geht es um die Frage wie:
 - Unter welchen Bedingungen ist eine Theorie testbar?
 - Wie kann man Theorien und Modelle optimal prüfen?
 - Wann wird eine Theorie verworfen?
 - Unter welchen Bedingungen gilt eine Theorie als bestätigt?

Für Popper steht die Frage der Testung und Rechtfertigung von Theorien im Zentrum der Wissenschaftstheorie.

Bevor wir uns dem nächsten Prinzip zuwenden, benötigen wir ein wichtiges Konzept



Konzept 2-4: Unbeschränkte und beschränkte Allaussagen:

Unter einer *unbeschränkte Allaussagen* versteht man eine Aussage, welche sich auf eine (potentiell) unendlich grosse Klasse von Entitäten bezieht (In der Psychologie handelt es sich hierbei meist um Menschen). Hier einige typische Beispiele für unbeschränkte Allaussagen:

- Alle Menschen verfügen über eine bestimmte Anzahl von angeborenen Grundemotionen, wie Wut, Furcht, Freude, etc.
- Vermummung und Verkleidung führen im Spiel von Kindern zur Erhöhung aggressiver Handlungen.
- Die Vorspiegelung falscher positiver Emotionen führt zu zunehmender sozialer Entfremdung.

In allen 3 Beispielen wird auf eine potentiell unendliche Klasse (Menschen, Kinder) Bezug genommen.

Beschränkte Allaussagen beziehen sich auf eine beschränkte Menge von Entitäten. Die Beschränkung erfolgt oft durch eine raum-zeitliche Spezifikation.

- Alle Personen, welche an diesem Experiment teilnahmen, erhielten Sfr. 20.-
- Alle Leute, die den Film sahen, waren nachher aufgewühlt.

In diesen Beispielen wird die Menge der betroffenen Personen durch eine zusätzliche Spezifikation eingeschränkt. Es handelt sich, da die Sätze die Form von Allsätzen aufweisen, um *ingeschränkte Allsätze*.



Bemerkung zur Wortwahl:

Manche Autoren machen eine Unterscheidung zwischen *Aussagen* und *Sätzen*, wobei erstere den Inhalt von Sätzen bezeichnen. In der nachfolgenden Darstellung wird – wie in der vorangegangenen – zwischen Sätzen und Aussagen kein Unterschied gemacht.

Das folgende Prinzip klärt die Beziehung zwischen All- und Existenzaussagen. Bei letzteren handelt es sich um Aussagen, welche die Existenz von etwas behaupten. Man nennt sie daher auch »Es gibt« Aussagen.



Prinzip 2-8: *Logische Beziehung zwischen All- und Existenzaussagen:*

Die logische Beziehung zwischen All- und Existenzaussagen lässt sich am einfachsten unter Verwendung einer symbolischen Notation demonstrieren.

Ein Allsatz wird formal wie folgt repräsentiert:

$$\forall x P(x) \quad (2-1)$$

In Worten:

»Auf alle Dinge, für welche die Grösse x stehen kann, trifft das Prädikat P zu.«

oder kürzer:

»Für alle x gilt P .«

Im Ausdruck (2-1) haben die Symbole folgende Bedeutung:

x ist eine *Individuenvariable*, welche über einzelne Objekte einer vorgegebenen Menge von Entitäten läuft, beispielsweise kann sich x auf die Elemente der Menge der Menschen beziehen.

P ist der Name eines *Prädikats* (einer Eigenschaft), z.B. Die Eigenschaft der Sterblichkeit.

\forall ist der *Allquantor*: Er drückt aus, dass der nachfolgende Ausdruck auf alle Objekte des Bereichs (auf alle Elemente der Menge) zutrifft.

Hier nun ein etwas komplexeres Beispiel einer Allaussage:

$$\forall x [M(x) \rightarrow S(x)] \quad (2-2)$$

In Worten:

»Für alle Lebewesen gilt: Wenn es sich um einen Menschen handelt, so ist sie sterblich«

oder kürzer:

»Alle Menschen sind sterblich«

Es gilt daher:

x ist eine Variable zur Bezeichnung der Lebewesen.

M bezeichnet die Eigenschaft des Menschseins.

S bezeichnet das Prädikat der Sterblichkeit.

Eine Existenzaussage hat demgegenüber die folgende formale Repräsentation:

$$\exists x P(x) \quad (2-3)$$

In Worten:

»Es gibt ein Ding x (aus der Zielmenge), auf welches das Prädikat P zutrifft.«

\exists bezeichnet den *Existenzquantor*.

Hier ein komplexeres Beispiel für einer Existenzaussage:

$$\neg \exists x \neg [M(x) \rightarrow S(x)] \quad (2-4)$$

In Worten:

»Es gibt kein Lebewesen, für das nicht gilt: Wenn es ein Mensch ist, so ist es sterblich«

oder kürzer:

»Es gibt keine Menschen, die nicht sterblich sind«

\neg bezeichnet die Negation des nachfolgenden Ausdrucks:

➤ $\neg \exists x P(x)$ bedeutet, dass es kein Objekt x (aus der betrachteten Zielmenge) gibt, auf welches das Prädikat P nicht zutrifft.

➤ $\neg \exists x \neg P(x)$ bedeutet daher, dass es kein Objekt x (aus der Zielmenge) gibt, auf welches das Prädikat P nicht zutrifft.

Es gelten nun die folgenden logischen Äquivalenzen zwischen All- und Existenzaussagen):

$$\begin{aligned}
& \forall x P(x) \Leftrightarrow \neg \exists x \neg P(x) \\
& \forall x \neg P(x) \Leftrightarrow \neg \exists x P(x) \\
& \neg \forall x P(x) \Leftrightarrow \exists x \neg P(x) \\
& \neg \forall x \neg P(x) \Leftrightarrow \exists x P(x)
\end{aligned}
\tag{2-5}$$

Die Äquivalenzen in (2-5) sind unmittelbar einsichtig. In Worten besagen sie:

- Das Prädikat P trifft auf alle Dinge x zu, gdw. (genau dann wenn) es kein Ding x gibt, auf welches P nicht zu trifft.
- Das Prädikat P trifft auf kein Ding x zu, gdw. es kein Ding x gibt, auf welches P zu trifft.
- Das Prädikat P trifft nicht auf alle Dinge x zu, gdw. es ein Ding x gibt, auf welches P nicht zu trifft.
- Das Prädikat P ist nicht auf alle Dinge x unzutreffend, gdw. es ein Ding x gibt, auf welches P zu trifft.

Die letzten beiden Äquivalenzen in (2-5) unterliegen der folgenden Einschränkung:

Die rechte Seite der Äquivalenz folgt aus der linken Seite nur, falls die Menge der Dinge, auf die sich die Variable x bezieht, nicht leer ist.

Die umgekehrte Folgerung (von rechts nach links) gilt immer. Die formale Repräsentation in (2-5) enthüllt ein Muster, nämlich:

- Ein Allsatz entspricht einem Existenzsatz mit Verneinung »vorne« und »hinten«: $\forall x \Leftrightarrow \neg \exists x \neg$.
- Aus der Tatsache, dass eine doppelte Verneinung einer Bejahung entspricht (d.h. zwei aufeinander folgende »¬« heben sich auf), ergeben sich die restlichen Äquivalenzen, z.B.: $\neg \forall x \Leftrightarrow \neg \neg \exists x \neg \Leftrightarrow \exists x \neg$.

Im Kontext der Prüfung von Theorien oder von unbeschränkten All-Aussagen haben die oben dargestellten logischen Äquivalenzen eine wichtige Bedeutung, weil sie das folgende Prinzip repräsentieren:

*Allaussagen implizieren nicht die Existenz von Objekten, aber sie **verbieten** die Existenz bestimmter Objekten.*

Wenn daher jemand behauptet, dass alle Menschen eine kleine Menge von Grundemotionen – wie Ärger, Furcht etc. – besitzen [$\forall x \text{GE}(x)$], so impliziert dies nicht die Existenz von Menschen. Die Aussage verbietet jedoch die Existenz von Menschen ohne derartige Basisemotionen. Daher steht sie zu einer Existenzaussage, welche behauptet, dass es einen Menschen ohne Grundemotionen gibt [$\exists x \neg \text{GE}(x)$], im Widerspruch, da ja gemäss (2-5) gilt: $\exists x \neg \text{GE}(x) \Rightarrow \neg \forall x \text{GE}(x)$, d.h. die Existenzaussage impliziert die Negation der Allaussage.

Diese Beziehung zwischen All- und Existenzaussagen bildet auch die Grundlage für das folgende wichtige Prinzip, welches für Poppers Konzeption von zentraler Bedeutung ist:



Prinzip 2-9: Asymmetrie in der Prüfung von unbeschränkten Allaussagen:

Unbeschränkte Allaussagen lassen sich grundsätzlich nicht verifizieren. Sie lassen sich jedoch im Allgemeinen falsifizieren.

- *Verifizieren* bedeutet, dass sich die Wahrheit einer Aussage feststellen lässt. Die Wahrheit einer unbeschränkten Allaussage lässt sich nicht feststellen, weil – aufgrund der potentiellen Unendlichkeit der Menge – nicht für alle Elemente geprüft werden kann, ob die Aussage zutrifft.
- *Falsifizieren* bedeutet, dass sich die Falschheit einer Aussage feststellen lässt. Die Falschheit lässt sich feststellen, wenn für ein Element der Menge die Falschheit gezeigt wurde.

Betrachtet man die oben als Beispiele angeführten unbeschränkten Allaussagen, so ist unmittelbar klar, dass diese sich zwar nicht verifizieren lassen. Eine Falsifikation ist jedoch prinzipiell möglich.

So ist es z.B. nicht möglich, die Allaussage über den Zusammenhang von Verkleidung und aggressiven Akten im Spiel von Kindern zu verifizieren, denn auch wenn bei allen untersuchten Kindern ein Zusammenhang auftrat, impliziert dies nicht die Wahrheit der Allaussage.

Die Aussage ist hingegen falsifizierbar, denn sobald ein Kind beobachtet wird, bei dem besagter Zusammenhang nicht auftritt, so ist die Aussage widerlegt.



Bemerkung zur Falsifikation unbeschränkter Allaussagen:

Oben wurde ausgesagt, dass unbeschränkte Allaussagen »im Allgemeinen« falsifizierbar sind. Dies bedeutet, dass es Ausnahmen gibt. Ein Beispiel hierfür ist die Aussage:

Alle Menschen sind sterblich.

Diese Aussage ist offensichtlich nicht falsifizierbar. Denn man müsste eine Person unendlich lange beobachten, um eine falsifizierende Instanz (d.h. einen unsterblichen Menschen) zu finden.

Das Problem liegt hier in der Prüfung des Prädikats *unsterblich* und nicht in der Tatsache, dass es sich um eine unbeschränkte Allaussage handelt.

Bevor wir uns Poppers deduktiv-falsifikationistischer Methode zuwenden benötigen wir noch die folgenden beiden Konzepte.



Konzept 2-5: Basissätze (Beobachtungssätze) und empirischer Gehalt (Teil 1):

Ein *Basissatz* (*Beobachtungssatz*) ist ein atomarer »Es gibt« - Satz mit Orts- und Zeitangabe, bzw. eine Menge derartiger atomarer Sätze, die mit »und« verknüpft sind.

Basissätze, welche im Widerspruch zu aus den Theorien abgeleiteten Sätzen stehen, sind *potentielle Falsifikatoren* der Theorie.

Eine Theorie, für welche Basissätze vorliegen, die im Widerspruch zu Ableitungen der Theorie stehen (und damit eine potentieller Falsifikatoren sind), besitzt einen *empirischen Gehalt*.



Bemerkung zur Widerspruchsfreiheit von Theorien:

Um Trivialitäten zu vermeiden, nehmen wir hier und im Folgenden an, dass Theorien widerspruchsfrei sind. Damit ist gemeint, dass aus einer Theorie nicht ein Satz und gleichzeitig sein Gegenteil abgeleitet werden kann.

Es ist einfach zu zeigen, dass aus einem widersprüchlichen Satzsystem jeder Satz abgeleitet werden kann (siehe z.B. Popper, 1989, Seite 58).

Bevor wir das Konzept des Basissatzes noch weiter durchleuchten, soll die folgende Frage beantwortet werden:



Frage:

Warum benötigt man, das Konzept eines Beobachtungssatzes überhaupt? Würde es nicht reichen, zu sagen, eine Beobachtung steht im Widerspruch zu einer Ableitung bzw. Vorhersage einer Theorie?

Die Antwort auf diese Frage lautet wie folgt: Beobachtungen können nicht im Widerspruch zu Sätzen stehen; nur Sätze können anderen Sätzen widersprechen. Daher verwendet man zur Prüfung von Theorien Sätze, die Beobachtungen repräsentieren, denn derartige Sätze können im Widerspruch zu Vorhersagen der Theorie stehen.



Bsp.2-18: Basissätze:

Gegeben: Die folgende Allaussage

Vermummung und Verkleidung führen im Spiel von Kindern zur Erhöhung aggressiver Handlungen.

Basissatz:

Es gibt ein Kind, das in dem am 10.6.1995 an der Universität Fribourg stattgefunden Experiment trotz Verkleidung beim Spielen keine erhöhte Anzahl aggressiver Handlungen zeigte.

Der Basissatz steht offensichtlich im Gegensatz zu der Allaussage, denn wenn alle Kinder bei Verkleidung eine Erhöhung des Anteils aggressiver Handlungen zeigen, so kann es keine Person geben, für die dies nicht gilt.

Die Allaussage hat daher einen empirischen Gehalt.

Der aufmerksamen Leserin wird sich an dieser Stelle vielleicht die folgende Frage aufdrängen:



Frage:

In der oben präsentierten Spezifikation von Basissätzen wurde angeführt, dass diese sich auf konkret beobachtete Dinge beziehen. Was bedeutet dies, wenn doch jede Beobachtung theoriegeladen ist?

In Prinzip 2-3 hatten wir Kriterien zur Beurteilung der Beobachtungsnähe von Aussagen und Begriffen behandelt. Hierbei wurde auch erwähnt, dass beobachtungsnaher Begriffe meist unproblematischer sind als beobachtungsfremde. Bei Basissätzen handelt es sich nun um Aussagen, welche so beobachtungsnah sind, dass sie für die Forschergruppe unproblematisch sind.



Konzept 2-6: Basissätze (Teil 2):

Basissätze sind Konventionen, die von der Forschungsgemeinschaft als unproblematisch akzeptiert werden.

Falls es keine Übereinstimmung mehr gibt bezüglich der Korrektheit von Basissätzen, so kommt es zum Stillstand des wissenschaftlichen Prozesses.

Popper schreibt hierzu:

»Es ist verständlich, daß sich auf diese Weise ein Verfahren ausbildet, bei solchen Sätzen stehenzubleiben, deren Nachprüfung „leicht“ ist, d.h. über deren Anerkennung und Verwerfung unter den verschiedenen Prüfern eine Einigung erzielt werden kann; [...] Sollte eines Tages zwischen wissenschaftlichen Beobachtern über Basissätze keine Einigung zu erzielen sein, so würde das bedeuten, daß die Sprache als intersubjektives Verständigungsmittel versagt. Durch eine solche Sprachverwirrung wäre die Tätigkeit des Forschers ad absurdum geführt; wir müssten unsere Arbeit am Turmbau der Wissenschaft einstellen.« (Popper, 1989; S. 70).

Man beachte den Zusammenhang dieser Konzeption mit Prinzip 2-4 und Prinzip 2-6, die beide auf die Notwendigkeit von Vorannahmen, die fürs erste akzeptiert werden, hinweisen.

Basissätze unterliegen – ebenso wie alle anderen Aussagen – der Kritik (Prinzip 2-5). Daher ist es möglich, dass auch Basissätze revidiert werden. Popper schreibt hierzu:

»Die Basissätze, bei denen wir jeweils stehenbleiben, bei denen wir uns befriedigt erklären, die wir als hinreichend geprüft anerkennen – sie haben wohl insofern den Charakter von Dogmen, als sie ihrerseits nicht weiter begründet werden. Aber diese Art von Dogmatismus ist harmlos, denn sie können ja, falls doch noch ein Bedürfnis danach auftreten sollte, weiter nachgeprüft werden.« (Popper, 1989; S. 70).



Bsp.2-19: Problematische Basissätze I: Ausfüllen von Fragebogen

Ein wichtiges Instrument in der Psychologie sind Fragebögen, welche Fragen mit mehreren Antwortoptionen enthalten. Die Antworten werden dann verwendet, um bestimmte Theorien des Untersuchers zu testen, z.B. Theorien über die Auswirkung des Führungsstils auf die Effizienz der Mitarbeiter.

Ein typischer Basissatz besagt dann, dass die Person p zum Zeitpunkt t am Ort o die Option X gewählt hat.

Der Basissatz ist völlig unproblematisch, solange man die gewählte Option eindeutig feststellen kann.

Nehmen wir nun an, die Versuchsperson hatte sich zuerst für eine andere Option als X entschieden, ihre Wahl jedoch später geändert. Hierbei habe sie die alte Markierung nur ungenügend entfernt, so dass diese noch sichtbar ist. In diesem Fall ist die Aussage des Basissatzes problematisch. Man wird folglich das Item entfernen.

Bsp.2-19 zeigt wie eine Beobachtung problematisch sein kann und damit auch der Basissatz, der sich auf diese Beobachtung bezieht. Das nächste Beispiel soll demonstrieren, dass Basissätze oft höhere Grade von Beobachtungsferne aufweisen können.



Bsp.2-20: Problematische Basissätze II: Kategorisierung von Items

Eine wichtige Methode in der Psychologie betrifft die Kategorisierung von Beobachtungsdaten – z.B. von verbalen Protokollen oder Videos. Hierbei teilen ein oder mehrere Rater Teilabschnitte des Protokolls jeweils einer aus einer Menge von fix vorgegebenen Kategorien zu.

Diese Methode käme vermutlich zum Einsatz, wenn man die oben gemachte Aussage über den Zusammenhang zwischen Verkleidung und Aggression im Spiel von Kindern (vgl. Bsp.2-18) testen möchte: Man filmt die Kinder beim Spiel und kategorisiert deren Handlungen in aggressive und nicht-aggressive.

Es ist offensichtlich, dass die kategorisierten Daten insofern problematisch sind, als die Urteile der Rater einfließen. Um diesen kritischen Aspekt zu minimieren, werden mehrere Rater verwendet und der Übereinstimmung (die so genannte *Interrater-Reliabilität*) wird gemessen. Falls diese einen ausreichenden hohen Grad erreicht, so werden die Daten als unproblematisch betrachtet.

Das letzte Beispiel zeigt, dass die Basissätze bzw. die Daten, auf die sie sich beziehen, eine gewisse Beobachtungsferne aufweisen können. Aber es kommt noch »dicker«! Es zeigt sich nämlich, dass praktisch alle Messungen in der Psychologie *theorieabhängig* sind. Damit sind natürlich auch die Basissätze, welche diese zum Inhalt haben, *theorieabhängig*.

Um dies zu verstehen benötigen wir zuerst das folgende Konzept:



Konzept 2-7: Theorieabhängige Messung:

Theorieabhängige Messungen verwenden eine Theorie bzw. ein theoretische Modell als Basis für die Messung.

Der Prozess der Messung setzt daher die Gültigkeit der zugrunde liegenden Theorie bzw. des Modells voraus.

Wir demonstrieren dieses Konzept anhand einiger einfacher Beispiele.



Bsp.2-21: Ein einfaches Prozessbaummodell zur Messung von Wissens- und Raterprozessen:

Gegeben:

Ein Menge von Testitems eines Leistungstests mit jeweils zwei Antwortmöglichkeiten, eine davon ist wahr, die andere falsch.

Zielsetzung:

Messung des Wissens der Person unter Berücksichtigung der Möglichkeit des Ratens einer korrekten Antwort.

Die grundlegenden Annahmen des Messmodells:

Das Messmodell in Abb. 2-3 macht folgende Annahmen:

1. Bei Präsentation eines Testitems ist der Proband mit Wahrscheinlichkeit π in der Lage, das zur Lösung relevante Wissen abzurufen (oberer Zweig in Abb. 2-3). In diesem Fall erfolgt die korrekte Antwort mit Wahrscheinlichkeit 1.
2. Mit Wahrscheinlichkeit $1-\pi$ ist das Wissen nicht abrufbar (entweder weil es nicht vorhanden ist oder weil nicht darauf zugegriffen werden kann). In diesem Fall rät die Person mit Wahrscheinlichkeit von 0.5 die korrekte Antwort.

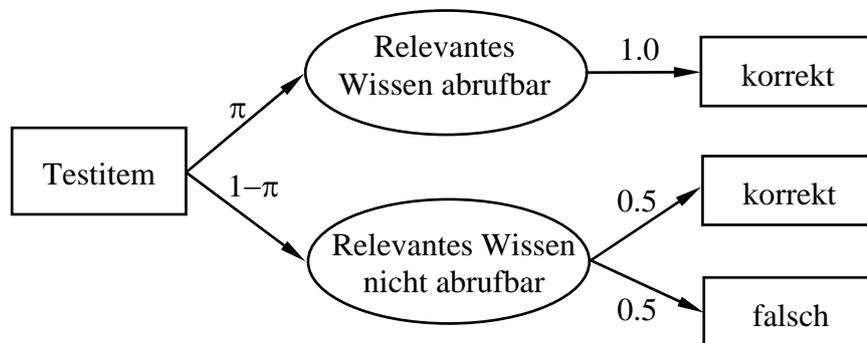


Abb. 2-3: Prozessbaum-Modell zur Messung von relevantem Wissen und Raten.

Bemerkung zur Grösse π :

Das Modell enthält die unbekannte Grösse π . Hierbei handelt es sich um einen statistischen Parameter des Modells. Statistische Parameter kennzeichnen so genannte Grundgesamtheiten oder Populationen. *Genauer:* Parameter kennzeichnen die Verteilung von Merkmalen einer Population. Sie werden daher auch *Kennwerte* genannt.

Die Werte von Parametern werden üblicherweise aus den Daten geschätzt. Im aktuellen Fall wird π mit Hilfe des Ausdrucks $2 \cdot p - 1$, geschätzt, wobei p der Anteil korrekter gelöster Items ist (falls $p < 0.5$, wird $\pi = 0$ gesetzt).

Im aktuellen Beispiel repräsentiert π die Wahrscheinlichkeit, dass eine Zielperson bei Vorliegen eines Testitems das relevante Wissen zur Lösung abrufen kann.

Die betrachtete Population besteht daher aus einem Pool von Testitems, die alle (annähernd) den gleichen Schwierigkeitsgrad aufweisen.

Beispielsweise könnte der Itempool aus eine Menge von multiple-choice Fragen mit je zwei Wahlmöglichkeiten zu einem Fachgebiet bestehen. Aus diesem Pool von Fragen, wird zufällig eine Stichprobe gewählt und dem Probanden vorgelegt.

Wie oben erwähnt, repräsentieren die Parameter den zentralen Aspekt einer Messung. Für das aktuelle Beispiel ist dies leicht nachzuvollziehen, denn der Prüfer ist daran interessiert, wie viel der Proband weiss. Der Anteil korrekter Antworten ist offensichtlich kein ideales Maß für das Wissen, da in dieser Grösse die Anteile von Wissen und Raten vermengt sind.

In diesem Zusammenhang spricht man auch von *reinen Prozessmaßen*. Hierbei handelt es sich um Maße welche einen Prozess oder ein Konstrukt »gereinigt« von anderen Einflüssen und Prozessen repräsentieren.

So kann man – falls das Modell korrekt ist – π als reines Maß des Wissens betrachten. Der Prozentsatz korrekter Antworten repräsentiert Wissen und Raten und ist hingegen kein reines Maß des Wissens der Person.

Bsp.2-21 demonstriert, wie ein einfaches Modell verwendet werden kann, um das Wissen einer Person zu messen. Eine korrekte Messung setzt voraus, dass das Modell zumindest annähernd korrekt ist. Dies ist jedoch in den meisten Fällen nicht der Fall, aus den folgenden Gründen:

- Die Testperson besitzt ein bestimmtes Wissen über das Testitem. Dies reicht zwar nicht aus, um mit Sicherheit eine korrekte Antwort zu geben (oberer Zweig im Modell von Abb. 2-3). Andererseits ist der Rateprozess nicht völlig zufällig (unterer Zweig im Modell von Abb. 2-3).
- Die Antwortoptionen sind derart konstruiert, dass sie die Testperson in die Irre führen sollen. Auch dies wird den Rateprozess in eine bestimmte Richtung lenken.

Da eine der beiden Bedingungen fast immer erfüllt ist, ist das Modell fast immer falsch.



Bemerkung:

Die Tatsache, dass das Modell falsch ist, hindert jedoch viele angewandte Tester nicht darin, die auf dem Modell beruhende Formel zur Ratekorrektur von Multiple-Choice Tests anzuwenden:

$$\hat{\pi} = \frac{n \cdot p - 1}{n - 1} \quad (2-6)$$

In Gleichung (2-6) symbolisiert p den Anteil korrekter Antworten, n die Anzahl der Antwortoptionen und $\hat{\pi}$ die (um das Raten korrigierte) Wahrscheinlichkeit einer korrekten Antwort (Das »Dach« über dem griechischen Buchstaben zeigt an, dass es sich um eine Schätzung handelt).

In vielen Fällen ist sich der Anwender nicht einmal bewusst, dass die angewendete Formel auf einem ungeeigneten Modell basiert.

Das Problem der Verwendung ungeeigneter Modelle innerhalb der Psychologie wird in Kapitel **xxxx** näher behandelt.

Man könnte nun meinen, dass es sich bei der in Bsp.2-21 diskutierten Situation um einen sehr speziellen Fall handelt, der in der Psychologie nur sehr begrenzt Anwendung findet. In Wirklichkeit ist praktisch jede Art der Messung in der Psychologie (wie auch in den anderen Wissenschaften) theorieabhängig. Der Grund hierfür liegt in der Tatsache, dass in der Psychologie durchwegs mentale Entitäten (Kapazitäten, Einstellungen, Traits, Stimmungen, etc.) gemessen werden. Alle diese Messungen basieren – gleichgültig, ob sich die Forscherin darüber im

Klaren ist oder nicht – auf Annahmen darüber, wie die mentalen Konstrukte mit den beobachteten Maßen in Verbindung stehen. Auch die Gütekriterien zur Beurteilung der Messungen (Validität und Reliabilität) sind theoretische Grössen, die nicht theorieunabhängig bestimmbar sind.

Die nächsten beiden Beispiele sollen dies verdeutlichen.



Bsp.2-22: Aktive Informationssuche bei Risikoentscheidungen

Die Forschungsgruppe um Oswald Huber untersucht die Bedeutung verschiedener Methoden zur Reduktion oder Elimination des Risikos bei Risikoentscheidungen. Hierbei handelt es sich um Entscheidungen, bei denen ein Verlust (relativ zum Status quo) möglich ist [Der Begriff *Verlust* meint im aktuellen Kontext nicht nur pekuniären Verlust, sondern jede Art von negativer Konsequenz, sei es Verminderung der Gesundheit, oder der Verlust eines Angehörigen].

Es zeigt sich, dass in manchen Fällen die Information über Möglichkeiten der Entschärfung des Risikos wichtiger ist als die Information über die Auftretenswahrscheinlichkeiten eines möglichen Verlusts (siehe z.B. Huber, Beutter, Montoya & Huber, 2001; Huber, Wider & Huber, 1997).

Um zu untersuchen, welche Informationen für die Entscheider relevant sind, wird die Methode der *Aktiven Informationssuche (AIS)* verwendet. Hierbei wird der Versuchsperson (Vp) zuerst eine Beschreibung der Entscheidungssituation gegeben. Sodann wird sie gebeten nach weiteren Informationen zu fragen, welche ihrer Meinung nach für eine gute Entscheidung benötigt werden.

Die Anwendung der Methode zur Untersuchung der Bedeutung von Risikoentschärfung basiert auf folgenden Annahmen:

- Die von den Vpn erfragte Information hat einen Einfluss auf deren Beurteilung der Entscheidungsalternativen.
- Die erfragte Information hat daher auch indirekt einen Einfluss auf die Entscheidungen.
- Information, die nicht erfragt wird, wird vom Entscheider als irrelevant betrachtet und beeinflusst daher das Entscheidungsverhalten nicht oder nur unwesentlich.

Diesen Annahmen liegt die Überlegung zugrunde, dass Personen nach genau jener Information fragen, die sie als relevant für ihre Entscheidung betrachten.

Diese Überlegung ist plausibel aber muss nicht notwendigerweise korrekt sein. So könnten die Vpn eine Art Brainstorming durchführen und alles Mögliche fragen, gleichgültig, ob es relevant ist oder nicht. Andererseits könnten sie gewisse Dinge nicht erfragen, weil sie bestimmte Vorannahmen haben und es daher für nicht notwendig halten, nach diesen Informationen zu fragen.

Bsp.2-22 zeigt, dass selbst die sinnvolle Verwendung von einfachen Daten, wie die von Personen erfragten Informationen, auf Annahmen beruht.



Bemerkung:

Die Frage, ob die Versuchsperson die Aufgabe und Situation in der gleichen Weise interpretiert wie der Experimentator ist zentral für die Interpretation vieler Experimente zu menschlichen Urteilen und Entscheidungen. In der Tat gab (und gibt) diese Problematik immer wieder Anlass zu ausgiebigen Diskussionen (siehe Kapitel xxx).

Zum Schluss sei noch ein etwas komplexeres Beispiel zur Problematik der theorieabhängigen Messung präsentiert.



Bsp.2-23: *Messung der Funktionen des Arbeitsgedächtnisses*

In den letzten Jahren wurde deutlich, dass die Konstrukte *Fluide Intelligenz* und *Arbeitsgedächtniskapazität* sehr eng verwandt sind (Ackerman, Beier & Boyle, 2005), wenn nicht sogar identisch (Kane, Hambrick, Tubolski, Wilhelm, Payne & Engle, 2004).

Um derartige Aussagen empirisch zu prüfen, müssen komplexe Messmodelle zur Messung der beiden Konstrukte verwendet werden. Im Folgenden wird ein Messmodell von Miyake, Friedman, Rettinger, Shah und Hegarty (2001) zur Messung verschiedener Funktionen des Arbeitsgedächtnisses vorgestellt. Es handelt sich hierbei um ein faktorenanalytisches Modell (vgl. Bsp.2-10).

Das Ziel der Untersuchung von Miyake et al. (2001) bestand in der Analyse des Verhältnisses von exekutiven und visuell-räumlichen Funktionen des Arbeitsgedächtnisses. Es wurden die folgenden 3 Faktoren (Konstrukte) postuliert:

- (1) *Visuell-räumlicher Verarbeitungsfaktor:* Betrifft die aktive Verarbeitung und Behaltensleistung von visueller Information.
- (2) *Visuelle Behaltensleistung:* Betrifft die reine Behaltensleistung von visuell-räumlicher Information.
- (3) *Exekutive Funktionen:* Funktionen der Planung, Steuerung und Konfliktresolution.

Zu jedem Faktor werden zwei Indikatoren (Aufgaben) definiert:

- (1) *Visuell-räumlicher Verarbeitungsfaktor:*
 - (i) Buchstabenrotation
 - (ii) Punktematrix: Verifikation einer Punktegleichung plus simultanes Behalten der Position eines Punktes in einem 5×5-Raster (vgl. Miyake et al., 2001, Seite 627).
- (2) *Visuelle Behaltensleistung:*
 - (i) Corsi-Blöcke: Eine präsentierte Sequenz von Punkten muss aus dem Gedächtnis wiedergegeben werden (vgl. Miyake et al., 2001, Seite 627).
 - (ii) Punktegedächtnis: Punkte in einem 5×5-Raster müssen aus dem Gedächtnis wiedergegeben werden (vgl. Miyake et al., 2001, Seite 627).
- (3) *Exekutive Funktionen:*
 - (i) Tower of Hanoi-Problem (3-7 Scheiben):
 - (ii) Zufallszahlengenerierung der Zahlen: 1, 2, ..., 9.

Das zugehörige Modell ist in Abb. 2-4 dargestellt.

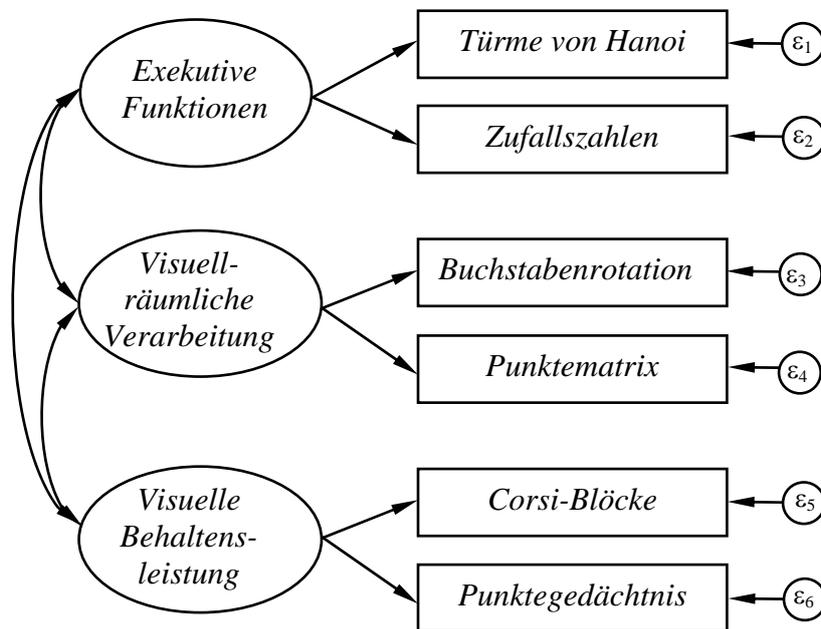


Abb. 2-4: Modell von Miyake et al. (2001) zu verschiedenen Funktionen des Arbeitsgedächtnisses.

Das Modell ermöglicht nun z.B. Aussagen über den Zusammenhang zwischen exekutiven Funktionen und visueller Verarbeitungskapazität zu untersuchen. Die Korrelation zwischen den beiden Faktoren, welche durch den Bogen mit den Doppelpfeilen repräsentiert ist, gibt darüber Auskunft.

Miyake et al. (2001) fanden, dass die beiden Faktoren *Exekutive Funktionen* und *Visuell-räumliche Verarbeitung* nicht zu unterscheiden sind, d.h. das zweifaktorielle Modell, mit einem Faktor, der sowohl die exekutive Funktionen wie die visuelle Verarbeitung repräsentiert (Abb. 2-5).

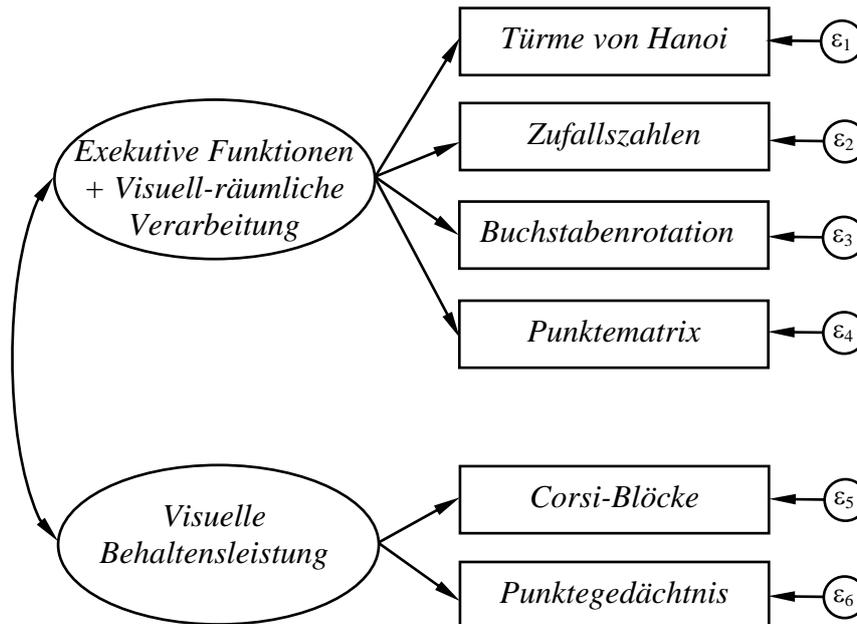


Abb. 2-5: Modifiziertes Modell von Miyake et al. (2001) zu verschiedenen Funktionen des Arbeitsgedächtnisses.

Der im aktuellen Kontext wichtige Aspekt besteht darin, dass die Beziehung zwischen den beiden Konstrukten nicht durch die Korrelation der direkt beobachteten Maße geprüft wird, sondern durch Verwendung von faktoranalytischen Messmodelle zur Messung der beiden Konstrukte. Die Korrelation zwischen den derart gemessenen Konstrukten wird ebenfalls mit Hilfe des Modells geschätzt.

Die Prüfung einer Theorie oder Gesetzhypothese, welche Aussagen über den Zusammenhang von Exekutiven Funktionen und visueller Verarbeitungskapazität macht, erfolgt daher bei Verwendung dieser Methode nur indirekt, nämlich unter Zuhilfenahme der als korrekt vorausgesetzten Messmodelle für die beiden Faktoren.

Die präsentierten Beispiele zeigen, dass die Beobachtungen in der Psychologie und folglich die Basissätze, welche sich auf diese beziehen, von unterschiedlicher Abstraktheit und Theorieladenheit sind. Dies bereitet jedoch keine Probleme, solange die Forschergemeinschaft die Theorien und Methoden als gültig akzeptiert und somit die darauf basierenden Beobachtungen als unproblematisch gelten.

Falls jedoch Zweifel an den Theorien und Modellen aufkommt und die gewonnenen Daten ihre Akzeptanz verlieren, so verlieren auch die Sätze, welche sich auf diese Daten beziehen, ihren Status als Basissätze. In diesem Falle muss entweder auf beobachtungsnähere Daten oder auf Beobachtungen, die auf verbesserten Modellen basieren, Bezug genommen werden.

Nach dieser ausführlichen Behandlung der Problematik theorieabhängigen Messung und des damit in Zusammenhang stehenden Problem des Findens einer Basis zur Prüfung von Theorien und Gesetzaussagen sei noch kurz auf einen anderen Aspekt von Basissätzen eingegangen.



Prinzip 2-10: *Basissätze sind keine Aussagen über Wahrnehmungserlebnisse:*

Als historische Vorläufer der Basissätze fungierten die so genannten *Protokollsätze*. Der zentrale Unterschied zu den Basissätzen bestand darin, dass letztere sich auf Wahrnehmungserlebnisse bezogen. Der zum Basissatzes:

Zum Zeitpunkt t befindet sich am Ort o ein Tisch.

äquivalente Protokollsatz könnte wie folgt lauten:

Ich, S. M., sehe zum Zeitpunkt t am Ort o einen Tisch.

Für Popper ist die Einbeziehung von Wahrnehmungserlebnissen ein Ausdruck eines unzulässigen *Psychologismus*, welcher seiner Meinung nach in der Wissenschaft nichts zu suchen hat. Dies hat folgenden Grund:

Die Schilderung eines subjektiven Erlebnisses führt in keiner Weise zu einer besseren Begründung von Basissätzen.

»Und was schließlich die psychologistische Basis betrifft, so ist es sicher richtig, daß der Beschluss, einen Basissatz anzuerkennen, sich mit ihm zu begnügen, mit Erlebnissen zusammenhängt – etwa mit Wahrnehmungserlebnissen; aber der Basissatz wird durch diese Erlebnisse nicht begründet; Erlebnisse können Entschlüsse, also auch Festsetzungen motivieren [vielleicht sogar entscheidend]; aber sie können einen Basissatz ebensowenig begründen wie ein Faustschlag auf den Tisch.« (Popper, 1989; S. 71).

Wir wenden uns nun der von Popper vorgeschlagenen Methode der Testung von Theorien zu.



Methode 2-2: *Deduktiv-falsifikationistische Methode der Prüfung von Theorien:*

Der Prozess der Prüfung wissenschaftlicher Theorien verläuft wie folgt:

1. Am Beginn stehen *Theorien*. Diese werden als gegeben vorausgesetzt. Deren Entstehung ist nicht Inhalt der Wissenschaftstheorie.

2. Aus den vorhandenen Theorien werden – unter Verwendung von Zusatzannahmen – *allgemeine Gesetze (Gesetzhypothesen)* abgeleitet.
3. Aus den allgemeinen Hypothesen werden – wiederum unter Verwendung von Zusatzannahmen – *konkrete Vorhersagen* deduziert.
4. Es wird geprüft, ob die Vorhersagen mit der Realität übereinstimmen. Dies bedeutet konkret, dass die Voraussagen mit den als unproblematisch betrachteten, mehr oder weniger theoriegeladenen, Beobachtungen verglichen werden.
5. Falls die Prüfung negativ ausfällt, so muss mindestens eine der beteiligten Theorien oder Zusatzannahmen falsch sein, da die Voraussagen logisch abgeleitet wurden. Dies folgt aus der logischen Korrektheit des *Modus tollens* (Vgl. Bsp. 2-1).
6. Fällt die Prüfung positiv aus, so gilt die Theorie als vorläufig bestätigt.

Man beachte, dass die Herleitung von Vorhersagen aus der Theorie Zusatzannahmen benötigt. Eine Herleitung von Vorhersagen ohne Zusatzannahmen ist nicht möglich.

Das folgende Beispiel exemplifiziert die einzelnen Schritte.



Bsp.2-24: *Prüfung einer Theorie: Abruf von Information aus dem Langzeitgedächtnis (LZG)*

1. *Die Theorie (Semantische Netzwerktheorie):*
 - Weltwissen wird im Gedächtnis in Form von elementaren Informationseinheiten und Beziehungen zwischen diesen gespeichert.
 - Der Abruf aus dem LZG beinhaltet zwei Prozesse:
 - (i) Einen *automatischen Prozess der Aktivierungsausbreitung* zwischen den Informationseinheiten über die semantischen Verbindungen. Dieser Prozess läuft sehr schnell ab und unterliegt nicht der bewussten Kontrolle der Person. Er kann daher auch nicht durch irgendwelche Strategien beeinflusst werden.
 - (ii) Einen langsamen kontrollierten *Prozess der Hemmung* irrelevanter Information. Dieser benötigt kognitive Ressourcen der Person. Dies bedeutet, dass die Person ihre Aufmerksamkeit auf den Prozess fokussieren muss. Wird letztere auf andere Inhalte gelenkt, so findet der Prozess der Hemmung nicht oder nur statt.

(iii) Der relative Grad der Aktivierung eines Inhaltes entscheidet darüber, wie schnell eine Person auf den gewünschten Inhalt zugreifen und damit eine Aufgabe, für deren Lösung der Inhalt benötigt wird, bewältigen kann.

Man beachte, dass nicht der absolute sondern der relative Grad der Aktivierung relevant ist. Sind nämlich viele Items hoch aktiviert, so können sie sich gegenseitig stören, was den Abruf erschwert. Nur wenn ein Item relativ zu den anderen hoch aktiviert ist, ist es gut verfügbar.

2. *Ableitung einer Gesetzeshypothese:*

Das Verhalten einer Person ist für abrufrelevanten Aufgaben bei kurzer Zeitdauer rein durch den automatischen Prozess der Aktivierungsausbreitung beeinflusst und bei längerer Zeitdauer durch den bewussten Prozess der Hemmung der automatischen Aktivierung.

Bemerkung:

Diese Gesetzeshypothese führt alleine noch zu keiner prüfbareren Vorhersage. Dies hat vor allem zwei Ursachen:

- Es muss einerseits konkretisiert werden, was unter »abrufrelevante Aufgaben« und unter einer kurzen und langen Zeitdauer zu verstehen ist. Ebenso muss festgelegt werden, worin die *Beeinflussung* des Verhaltens besteht.
- Die theoretischen Konstrukte der automatischen Aktivierungsausbreitung und der Hemmung müssen messbar gemacht werden.

Die Durchführung dieser Aufgaben lassen sich unter dem Begriff der *Operationalisierung* subsumieren.

3. *Operationalisierung der Gesetzesaussage (Das Experiment von Neely, 1977):*

Die abrufrelevante Aufgabe war eine lexikalische Entscheidungsaufgabe. Sie bestand darin, für ein präsentiertes Item möglichst schnell zu entscheiden, ob es sich um ein Wort der Sprache handelt oder um ein Nichtwort (Letzteres wurde aus einem Wort durch Änderung eines Buchstaben erzeugt).

➤ *Ablauf des Experiments:*

- (i) Zuerst wird ein so genanntes Primewort am Bildschirm präsentiert, z.B.: »Vogel«, »Körper«.
- (ii) Anschliessend folgte die Darbietung des Zielwortes (am Bildschirm), z.B. »Spatz«, »Herz« (Nichtworte: »Sputz«, »Hers«).

- (iii) Der zeitliche Abstand zwischen Präsentation des Prime- und Zielwortes wurde variiert. Unter anderem wurden das Zielwort entweder 250 ms oder 1000 ms nach dem Primewort präsentiert.
- (iv) Sobald das Zielwort präsentiert worden war, sollte die Vp möglichst schnell durch Drücken der entsprechenden Taste angeben, ob es sich um ein Wort oder Nichtwort handelte. Die RZ der Vp war das entscheidende Maß.

➤ *Die Struktur des Stimulusmaterials:*

Tab. 2-1 illustriert die Struktur des verwendeten Stimulusmaterials.

In der Bedingung *erwartet* folgt dem Prime in 2/3 der Fälle als Zielwort ein zum Primewort semantisch verwandtes Wort (der Name eines Vogels) und in 1/3 der Fälle ein nicht verwandtes (z.B. der Name eines Körperteils oder der Name eines Gebäudeteils).

In der Bedingung *unerwartet* folgt dem Prime in 2/3 der Fälle als Zielwort ein zum Primewort unverwandtes Wort (z.B. der Name eines Vogels) und nur in 1/6 der Fälle ein semantisch verwandtes (ein Körperteil).

<i>Bedingung</i>	<i>Prime</i>	<i>Zielwort</i>	<i>p</i>
Erwartet	»Vogel«	»Spatz«	2/3
		»Herz«	1/6
		»Fenster«	1/6
Unerwartet	»Körper«	»Spatz«	2/3
		»Herz«	1/6
		»Fenster«	1/6

Tab. 2-1: *Struktur des Stimulusmaterials im Experiment von Neely (1977) [p = Auftretenswahrscheinlichkeit].*

➤ *Die der experimentellen Manipulation zugrunde liegende Idee:*

Die Vp weiss, dass in der unerwarteten Bedingung in den meisten Fällen ein semantisch nicht verwandtes Wort – nämlich eine Vogelart – auf den Prime (Körper) folgen wird.

Um daher ihre Reaktionszeit zu minimieren, sollte sie den Prime dazu verwenden, um sich innerlich auf die Kategorie aus der das Zielwort vermutlich stammt (Körperteil) einzustellen.

4. Vorhersage:

- Wird das Zielwort 250 ms nach dem Prime präsentiert, so zeigt sich sowohl in der erwarteten wie in der unerwarteten Bedingung ein Erleichterungseffekt für die mit dem Prime semantisch verwandten Zielworte (gegenüber unverwandten Worten). Das bedeutet, dass die RZ zur Durchführung der lexikalischen Entscheidung für semantisch verwandte Worte verkürzt ist.
- Wird das Zielwort 1000 ms nach dem Prime präsentiert, so zeigt sich in der *erwarteten* Bedingung ein Erleichterungseffekt für die semantisch verwandten Zielworte.
In der *unerwarteten* Bedingung zeigt sich hingegen ein Erleichterungseffekt für die Worte aus der wahrscheinlicheren (semantisch unverwandten) Kategorie.

5. Interpretation:

- Der Erleichterungseffekt bei 250 ms für semantisch verwandte Worte in beiden Bedingungen zeigt, dass die semantisch verwandten Items automatisch aktiviert werden. Die Vp kann sich dem nicht entziehen.
- Der Erleichterungseffekt bei 1000 ms für semantisch nicht verwandte Worte in der unerwarteten Bedingung zeigt, dass die Vp nun ausreichend Zeit hatte, sich auf die Kategorie, welche wahrscheinlich auftreten wird, zu konzentrieren.

6. Abschliessende Bemerkung:

Die Darstellung der Operationalisierung sollte klar machen, wie die oben im Zusammenhang mit der Formulierung der Gesetzeshypothese beschriebenen Aufgaben gelöst wurden:

- Als abrufrelevante Aufgabe wurde eine lexikalische Entscheidungsaufgabe gewählt. Diese ist insofern geeignet, als sie den Zugriff auf das semantische Gedächtnis benötigt.
- Die lange und kurze Zeitdauer entspricht den 1000 ms bzw. 250 ms zwischen der Präsentation von Prime- und Zielwort.
- Die Beeinflussung ist im aktuellen Fall ein Erleichterungseffekt.

- Die automatische Aktivierung wird durch die Erleichterung bei einer kleinen Zeitdifferenz zwischen Prime- und dem *semantisch verwandten* Zielitem gemessen. Der automatische Aspekt der Aktivierung zeigt sich darin, dass ein Erleichterungseffekt auch in der *unerwarteten* Bedingung auftritt, in der eine Umorientierung auf die zu erwartete Kategorie strategisch sinnvoll ist. Die Person ist daher scheinbar nicht in der Lage, die Aktivierung (der in diesem Fall nicht wenig hilfreichen Information) zu kontrollieren.
- Die Hemmung wird gemessen durch den Erleichterungseffekt bei einer grossen Zeitdifferenz zwischen dem Prime und dem *semantisch nicht verwandten* Zielitem in der *unerwarteten* Bedingung. Dies deutet darauf hin, dass die Person die automatisch aktivierten Informationseinheiten unterdrückt hat, und daher schneller auf die relevante Information zugreifen kann.

Die gegebene Darstellung, welche nur die wesentlichen Aspekte des Experiments von Neely (1977) präsentierte, sollte den komplexen Prozess der Ableitung konkreter Vorhersagen aus theoretischen Annahmen verdeutlichen.

Die Folgerungen aus dem Experiment hängen vom Ergebnis ab:

1. Die Vorhersagen werden durch die Daten bestätigt (was tatsächlich zutrifft): In diesem Fall gilt die Theorie als *bestätigt* bzw. als *nicht falsifiziert*. Dies bedeutet jedoch keineswegs, dass die Theorie in allen Aspekten gut gesichert ist. Ein positives Ergebnis legt z.B. nicht zwingend nahe, dass die Aktivierung in der unerwarteten Bedingung tatsächlich unterdrückt wird. Man kann den Erleichterungseffekt für semantisch unverwandte Worte in dieser Bedingung auch dahingehend interpretieren, dass die Vp während der 1000 ms intern einen neuen Prime, welcher semantisch mit den Items aus der wahrscheinlichsten Kategorie (Vogel) verwandt ist, generiert. Dies führt zu einer automatischen Aktivierung der Elemente dieser Kategorie. Eine *aktive Suppression* der ursprünglich aktivierten Konzepte ist nicht unbedingt notwendig, wenn man davon ausgeht, dass die Aktivierung relativ schnell von selbst abnimmt.



Bemerkung:

Bei den Begriffen *Aktivierungsausbreitung* und *Hemmung* handelt es sich um theoretische Metaphern. Die Bedeutung derartiger Metaphern in der psychologischen Theoriebildung wird in Kapitel xxx behandelt.

2. Die Vorhersagen werden durch die Daten nicht bestätigt: In diesem Fall ist klar, dass mindestens eine der Annahmen in der Ableitung nicht korrekt sein kann. Da die Ableitung der Vorhersagen auf einem komplizierten Geflecht von Annahmen beruht, ist nicht von

vornherein klar, worauf das Scheitern zurückzuführen ist. So könnte es möglich sein, dass das verwendete Material ungeeignet war, um die Theorie zu testen, indem z.B. die Prime- und Zielworte nicht ausreichend semantisch verwandt waren. In diesem Falle müssen weitere Experimente durchgeführt werden, um genauer zu bestimmen, was die Ursache für das Scheitern des Experiments darstellte.



Bemerkung:

Diese Problematik, welche oft als Kritik an Poppers Falsifikationismus verwendet wird (m. E. zu Unrecht, da Popper dieses Problem durchaus erkannt hat [z.B. in Popper, 1974 (Angabe präzisieren xxx)]), wird im nächsten Kapitel eingehender besprochen.

Wir betrachten als Nächstes einen Spezialfall der Prüfung von Theorien, dem Popper grosse Bedeutung einräumt.

2.2.3.1 EXPERIMENTUM CRUCIS

Eine wichtige Methode der Prüfung von Theorien bildet die Durchführung eines so genannten *Experimentum crucis*. In diesem Fall werden zwei Theorien geprüft, welche einander widersprechende Vorhersagen machen. Die Verwendung eines Experimentum crucis zum Vergleich von zwei Theorien ist in der Psychologie weit verbreitet. Die folgenden beiden Beispiele demonstrieren das Prinzip.



Bsp.2-25: *Experimentum crucis I: Das klassische Experiment von Rock (1957):*

(i) *Einleitung.*

Das Experiment richtet sich gegen die Annahme der behavioristischen Lerntheorien, wonach beim Lernen Assoziationen inkrementell (d.h. in kleine Schritten) durch Wiederholung ausgebildet werden.

(ii) *Die Problemstellung:*

Beim Lernen einer Liste von Itempaaren gibt es mindestens zwei alternative Möglichkeiten, wie die wiederholte Darbietung des Lernmaterials zu einer Leistungsverbesserung führen kann:

- (a) Die Assoziation zwischen den beiden Items jeden Itempaars auf der Liste wird graduell in jedem Durchgang gestärkt.
- (b) In einem Durchgang werden die Assoziationen für nur für einen Teil der Itempaare ausgebildet, während für die restlichen Paare keine Assoziationen gebildet werden. Somit erfolgt die Bildung nicht graduell, sondern es liegt ein Alles-oder-Nichts – Prozess vor.

(iii) *Bemerkung zum Paarassoziationslernen:*

Beim Paarassoziationslernen wird immer ein Item (das Stimulus-Item) eines Paares präsentiert und die Vp antwortet mit dem zweiten Item (dem Response-Item) des Paares.

(iv) *Die Methode der Ersetzung ungelernter Items:*

Um diese beiden Alternativen gegeneinander zu testen, wurde die Leistung zweier Gruppen verglichen:

(a) *Kontrollgruppe:*

Die Vpn lernten die Items einer Liste, solange bis die gesamte Liste korrekt wiedergegeben werden konnte.

(b) *Experimentalgruppe:*

Jedes Response-Item, welches in einem Durchgang nicht korrekt wiedergegeben werden konnte, wurde durch ein neues ersetzt (d.h. das zugehörige Stimulus-Item wurde mit einem neuen Response-Item gepaart). Diese Ersetzung eines nicht reproduzierten Response-Items erfolgte auch dann, wenn das Item in einem früheren (als dem letzten) Durchgang korrekt wiedergegeben worden war. Die Ersetzung erfolgte durch Zufallsauswahl aus einem zuvor festgelegten Pool von Items, aus dem auch die Response-Items der Anfangspaare zufällig gewählt worden waren.

(v) *Hypothese / Argument:*

Falls die Assoziation zwischen Stimulus und Response inkrementell gebildet wird, sollten die Vpn der Kontrollgruppe einen deutlichen Vorteil besitzen und schneller lernen (Vorhersage der assoziativen Lerntheorien).

Wird hingegen eine Assoziation zwischen den Items eines Paares entweder gelernt oder nicht, so sollte es keinen Unterschied zwischen den Gruppen geben (Vorhersage der Alles-oder-Nichts Hypothese).

(vi) *Ergebnisse:*

In beiden durchgeführten Experimenten (Experiment 1 verwendete Buchstaben-Zahlen – Paare, Experiment 2 sinnlose Silben) gab es nur minimale Unterschiede zwischen den Gruppen:

Experiment 1:

➤ Durchschnittliche Anzahl Durchgänge (Dgn) bis zur Erreichung des Lernziels:

Kontrollgruppe:	4.55
Experimentalgruppe:	4.35

➤ Durchschnittliche Anzahl Fehler bis zur Erreichung des Lernziels:

Kontrollgruppe:	17.9
Experimentalgruppe	17.6

Experiment 2:

➤ Durchschnittliche Anzahl Dgn bis zur Erreichung des Lernziels:

Kontrollgruppe:	8.1
Experimentalgruppe	8.1

➤ Durchschnittliche Anzahl Fehler bis zur Erreichung des Lernziels:

Kontrollgruppe:	26.7
Experimentalgruppe	29.2

(vii) *Interpretation:*

Die Ergebnisse sprechen eindeutig für einen Prozess des Alles-oder-Nichts-Lernens und gegen einen inkrementellen Erwerb.

Diese Befunde wurden unter verschiedenen Bedingungen wiederholt. So stellte sich z.B. heraus dass auch die Konditionierung des Lidschlagreflexes ein Alles-oder-Nichts Phänomen ist, wobei der irrige Eindruck einer graduellen Ausbildung des Reflexes durch Zusammenfassung der Daten verschiedener Individuen entsteht (Estes, 1960).

Hier noch eine zweite Untersuchung, welche als Realisation eines Experimentum crucis betrachtet werden kann.



Bsp.2-26: *Experimentum crucis II: Lernen von Kausalzusammenhängen*

(i) *Einleitung:*

Das folgende Experiment zielt wiederum gegen einfache assoziative Lerntheorien, welche das Lernen von Kausalzusammenhängen als Bildung von Assoziationen zwischen Ursache und Wirkung verstehen.

(ii) *Theoretischer Hintergrund:*

Beim kausalen Lernen besteht die Aufgabe darin herauszufinden, welche von mehreren möglichen Ereignissen Ursache für ein Zielereignis bilden. Der Konkretheit halber nehmen wir an, die Aufgabe bestünde darin herauszufinden, welche der beiden Lebensmittel *Eier* oder *Milch* bei einer bestimmten Person eine allergische Reaktion auslöst.

Hierbei werden der Vp nacheinander einzelner Fälle präsentiert, in welcher angegeben ist, welche Ursachen vorhanden waren und ob der Effekt vorlag oder nicht, z.B.

Eier → Allergie
 Eier + Milch → Allergie
 ---- → Keine Allergie

Im ersten Fall hatte die Person Eier gegessen und eine Allergie ausgebildet. Im zweiten Fall hatte sie Eier und Milch zu sich genommen und eine allergische Reaktion gezeigt. Im dritten Fall hatte sie weder Eier noch Milch zu sich genommen und keine Allergie ausgebildet.

Assoziative Lerntheorien nehmen nun an, dass im Lernprozess zwischen den möglichen Ursachen und der Wirkung Assoziationen gebildet werden. Je stärker eine Assoziation zwischen möglicher Ursache und Wirkung, desto eher wird erstere als Ursache angesehen (siehe).

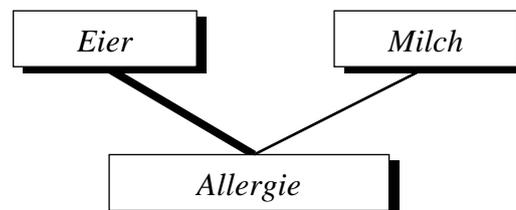


Abb. 2-6: Assoziatives Modell des Lernens von Ursache-Wirkungs-Zusammenhänge: die Dicke der Linien zwischen den Ursachen und der Wirkung repräsentiert den Grad des erlernten kausalen Zusammenhanges.

Assoziative Lerntheorien sind in der Lage, ein wichtiges Phänomen im Zusammenhang mit Kausallernen zu erklären, das Phänomen der *Blockierung* (vgl. Bsp.2-33).

Blockierung tritt immer dann auf, wenn im Lernprozess eine neue mögliche Ursache immer nur zusammen mit einem anderen Ereignis auftritt, von dem bereits bekannt ist, dass sie eine Wirkung besitzt. In diesem Fall kann die kausale Wirkung der neuen möglichen Ursache nicht eindeutig geklärt werden und die Personen neigen dann eher dazu, dass sie diese als nicht wirksam betrachten.



Bemerkungen zur Blockierung:

1. Das Phänomen der Blockierung ist eine Variante des Phänomens des *Discountings* (Kelley, 1972), wonach ein Ereignis als mögliche Ursache abgewertet wird, wenn bekannt wird, dass gleichzeitig mit dem Ereignis ein anderer Faktor anwesend war, von dem bekannt ist, dass er zum beobachteten Effekt führt.

Dieses Prinzip ist im gerichtlichen Kontext von grosser Bedeutung, wie das folgende reale Beispiel zeigt:

Am 27. Juni 1996 ereignete sich im Inselspital in Bern ein tragischer Vorfall: Im Zuge einer Operation am offenen Herzen wurde dem Patienten R eine Kanüle mit zwei Ausgängen in die rechte Halsschlagader gesetzt. Diese sollte zur Unterstützung des Blutkreislaufes während der Operation dienen.

Einer der beiden Ausgänge war jedoch verstopft, so dass die Infusion nicht ins Blut gelangte, sondern in das umliegende Gewebe. Dies wurde vom medizinischen Personal nicht sofort bemerkt, da der Hals des Patienten durch Tücher verdeckt war. In der Folge kam es zu einer Unterversorgung des Gehirns mit Sauerstoff, was bei R zu schweren geistigen Schäden führte.

R lebt heute in einem Pflegeheim für Schwerbehinderte. Er hat keinen gesetzlichen Anspruch auf Schadenersatz, da ein eindeutiger Kausalzusammenhang zwischen der fehlerhaften Behandlung und der Unterversorgung des Gehirns mit Sauerstoff nicht bewiesen werden kann.

Das Krankenhaus bestreitet zwar nicht einen Kunstfehler begangen zu haben, argumentiert jedoch, dass die Unterversorgung ebenso durch die Verstopfung eines Gefäßes aufgrund einer Fettablagerung während der Operation verursacht worden sein könnte.

Das Krankenhaus konnte sich auf diese Weise seiner Verantwortung entziehen, indem es eine andere Ursache fingierte, die den Effekt auslösen kann.

2. Das Phänomen der Blockierung widerspricht älteren Theorien des assoziativen Lernens, wonach es hinreichend für die Bildung einer Assoziation zwischen 2 Stimuli ist, dass diese nur oft genug zusammen auftreten.

Kehren wir nun wieder zu Bsp.2-26 zurück.

(iii) *Die Theorie der kausalen Strukturen:*

Waldmann und Holyoak (1992) nahmen an, dass Personen abstraktes Wissen über Kausalstrukturen besitzen, welche den kausalen Lernprozess mit beeinflussen.

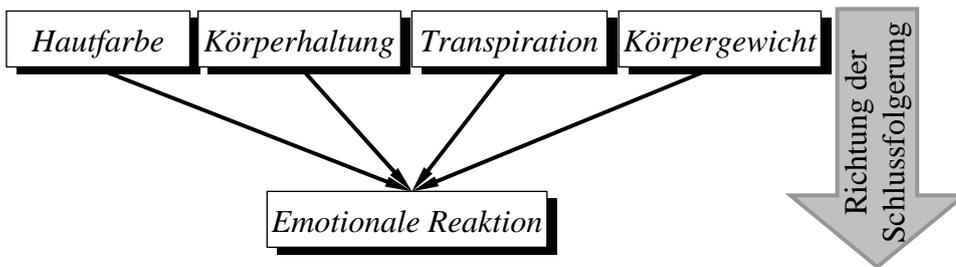
Konkret verwendeten sie die folgenden beiden Kausalstrukturen:

Konkret verglichen sie das Lernen bei Vorliegen der folgenden beiden Kausalstrukturen (Abb. 2-7):

- *Struktur mit multiplen Ursachen:* Mehrere Ursachen haben die gleiche Wirkung.
- *Struktur mit multiplen Wirkungen:* Eine einzige Ursache hat mehrere Wirkungen.

Waldmann und Holyoak (1992) verwendeten zwei Szenarien derart, dass die Ursachen in einem Szenarium identisch zu den Wirkungen im zweiten Szenarium war (Abb. 2-7):

Kausalstruktur der multiplen Ursachen:



Kausalstruktur der multiplen Wirkungen:

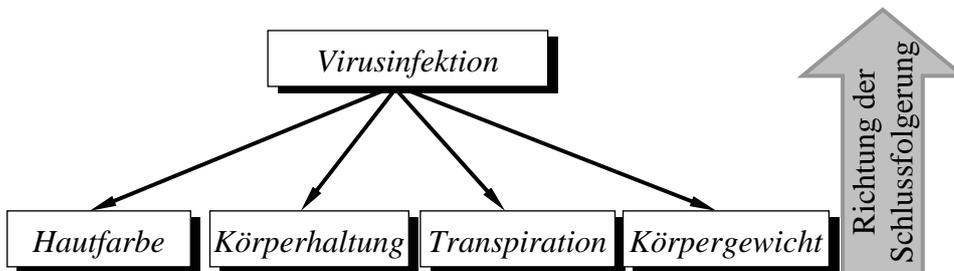


Abb. 2-7: Kausalstrukturen und Schlussrichtungen in der Bedingung mit multiplen Ursachen bzw. mit multiplen Wirkungen in Experiment 1 von Waldmann und Holyoak (1992).

Den Vpn der Gruppe mit *multiplen Ursachen* wurde mitgeteilt, dass bestimmte körperliche Charakteristika die emotionale Reaktion eines Beobachters beeinflussen. Dieser Einfluss sei nicht direkt beobachtbar, könne aber mittels physiologischer Tests gemessen werden.

Die Aufgabe bestand darin herauszufinden, ob eine bestimmte körperliche Charakteristik eine Wirkung auf die emotionale Reaktion eines Beobachters hat.

Den Vpn der Gruppe mit *multiplen Wirkungen* wurde mitgeteilt, dass eine Virusinfektion bestimmte körperliche Zustände verursacht.

Die Aufgabe bestand darin herauszufinden, welche Symptome für das Vorliegen eines Virus indikativ sind.

Die Vpn erhielten also in beiden Gruppen exakt die gleichen Cues, d.h. Hinweise, mit deren Hilfe auf ein anderes Ereignis (Emotionale Reaktion oder Virusinfektion) geschlossen werden musste.

Der zentrale Unterschied zwischen den beiden Gruppen bestand daher einzig in der Kausalstruktur und in der Richtung in die geschlossen werden musste: Entweder in Kausalrichtung (bei der Struktur mit multiplen Ursachen) oder entgegen die Kausalrichtung (bei der Struktur mit multiplen Wirkungen)

(iv) *Vorhersagen der verschiedenen Theorien:*

Die assoziative Lerntheorie, welche annimmt, dass kausale Beziehungen als ungerichtet Assoziationen zwischen Ursache und Wirkung repräsentiert werden, sagt vorher, dass die beiden Bedingungen exakt zum gleichen Ergebnis führen, da die Cues in beiden Bedingungen exakt identisch waren.

Die Theorie der Kausalmodelle, welche annimmt, dass die die Kausalstrukturen das Lernen von Ursache-Wirkungs-Beziehungen beeinflusst, sagt einen Unterschied zwischen den beiden Gruppen vorher.

Im Speziellen sollte in der *Bedingung mit multiplen Ursachen* eine Blockierung auftreten, falls ein Cue redundant ist, d.h. immer nur zusammen mit einem anderen Cue auftritt, der bereits als wirksame Ursache etabliert wurde.

Tritt beispielsweise in der Struktur in Abb. 2-7: eine bestimmte Hautfarbe (blass) immer nur mit einer hohen Transpiration auf und wurde durch vorangegangenes Lernen bereits die Tatsache etabliert, dass hohe Transpiration beim Beobachter eine emotionale Reaktion hervorruft, dann sollte die Hautfarbe als eher irrelevant für das Auftreten einer emotionalen Reaktion betrachtet werden, da deren Auftreten ja bereits durch das Vorhandensein einer Ursache (Transpiration) erklärt wird.

In der *Bedingung mit multiplen Wirkungen* sollte hingegen keine Blockierung gefunden werden, denn für Wirkungen gibt es kein zur Blockierung äquivalentes Prinzip:

Ist z.B. bekannt, dass eine Virusinfektion zu einer erhöhten Transpiration führt, so schliesst dies in keiner Weise andere Wirkungen – wie eine blasse Hautfarbe – aus.

(v) *Ergebnisse:*

Wie zu erwarten, bestätigen die Ergebnisse die Theorie von Waldmann und Holyoak (1992).

Ähnliche Ergebnisse wurden nicht nur in zahlreichen weiteren Studien an Erwachsenen gefunden, sondern auch bei Kindern (Gopnik, Glymour, Sobel, Schultz, & Kushnir, 2004).

Mit diesen Beispielen beschliessen wir die Darstellung der Popperianischen Konzeption der Psychologie als Wissenschaft und wenden uns nachpopperianischen Konzeptionen zu.

2.3 Konzeptionen der Wissenschaft nach Popper

In diesem Abschnitt werden Konzeptionen, welche nach Popper und zum Teil als Kritik an Poppers Konzeption entstanden sind, behandelt. Die zentrale Kritik an Popper betrifft dessen Falsifikationismus. Wir werden uns daher im ersten Teil verschiedenen Argumenten und Wissenschaftskonzeptionen zuwenden, welche im Zusammenhang mit der Kritik an Poppers falsifikationistischer Position stehen.

Popper wird oft ein strenger Falsifikationismus vorgeworfen. Dieser beschreibe weder die Vorgehensweise von Wissenschaftlern korrekt noch sei er sinnvoll für den Fortschritt der Wissenschaft. Dieser Vorwurf ist m. E. nicht gerechtfertigt. Denn einerseits ist es für Popper durchaus sinnvoll, eine Theorie, deren Vorhersagen nicht bestätigt wurden, weiter beizubehalten, indem so genannten *ad hoc Annahmen* zur Stützung der Theorie getroffen werden. Diese sollten jedoch den empirischen Gehalt und somit der Prüfbarkeit der Theorie erhöhen.



Bemerkung zu ad hoc Annahmen:

Das wohl beste Beispiel für eine erfolgreiche ad hoc Annahme bietet die Entdeckung des Planeten Neptun. Nachdem festgestellt worden war, dass die Bahn des Uranus nicht den Kepler'schen Gesetzen entsprach, wurde ein neuer Planet postuliert, dessen Bahn von John Adams 1843 und unabhängig (und exakter) von Urbain Le Verrier 1846 berechnet wurde.

Der neue Planet Neptun wurde schliesslich 1846 von Johann Gottfried Galle entdeckt.

Ad hoc Annahmen und Alternativerklärungen spielen in der Psychologie eine wichtige Rolle. Sie werden ausführlich in Abschnitt 4.3 behandelt.

Es muss auch angemerkt werden, dass für Popper nicht-falsifizierbare Theorien durchaus wertvoll sein können. So betrachtet Popper die Darwinsche Evolutionstheorie als nicht falsifizierbar, hält sie aber sehr wohl für eine wichtige Theorie.

2.3.1 Zwei grundlegende Probleme: Holismus und die Unterbestimmtheit von Theorien

Welche Gründe sprechen nun gegen Poppers Idee der Falsifikation als zentralen methodischen Ansatz zur Prüfung von Theorien? Im Folgenden werden zwei Probleme besprochen, welche als mögliche Einwände gegen die Falsifikation von Theorien vorgebracht wurden: Holismus und die Unterbestimmtheit von Theorien durch Daten.

2.3.1.1 DAS PROBLEM DES HOLISMUS

Das Problem des Holismus wurde erstmals von Pierre Duhem (1861–1916) klar formuliert.



Konzept 2-8: *Problem des Holismus im Kontexts der Prüfung von Theorien:*

Theorien können nicht isoliert getestet werden. Denn für die Herleitung von Vorhersagen werden Zusatzannahmen benötigt. Daher kann immer nur der Gesamtkomplex bestehend aus Theorie und Zusatzannahmen geprüft werden.

Fällt die Prüfung daher negativ aus, so bleibt unklar, ob die Ursache für die Falsifikation in einer fehlerhaften Theorie oder in falschen Zusatzannahmen liegt.

Das Problem des Holismus läuft darauf hinaus, dass auch die Falsifikation von Theorien niemals als endgültig betrachtet werden kann. Es ergibt sich daher das folgende Prinzip:



Prinzip 2-11: *Vorläufigkeit der Falsifikation von Theorien:*

Theorien können immer nur vorläufig und nie endgültig falsifiziert werden.

Das Prinzip lässt sich einfach illustrieren. Kehren wir hierzu nochmals zu Bsp.2-24 zurück:



Bsp.2-27: *Abruf von Information aus dem Langzeitgedächtnis (Fortsetzung von Bsp.2-24)*

Angenommen, die Vorhersagen von Neely (1977) hätten sich nicht bestätigt. Im Speziellen sei angenommen, dass das Ergebnis wie folgt aussah:

Die Präsentation eines semantisch verwandten Primewortes führt zu keiner kurzfristigen Erleichterung der lexikalischen Aufgabe (Diskrimination zwischen Worten und Nonworten).

Das Nichteintreten des gewünschten Ergebnisses widerlegt nicht direkt die Theorie der automatischen Aktivierung von semantisch verwandten Worten im Langzeitgedächtnis, sondern nur *die Gesamtheit aus Theorie und Zusatzannahmen*. Zusatzannahmen sind z. B.:

- Die gewählten verwandten Zielworte sind mit den Primeworten semantisch ausreichend verwandt, so dass eine messbare Aktivierung durch die Primeworte stattfindet.
- Das Intervall von 250 ms ist ausreichend lang, damit die Aktivierungsausbreitung stattfinden kann und kurz genug, dass strategische Prozesse ausgeschlossen werden können.

Bemerkung:

Das kritische Problem stellt eher die lange Dauer des Intervalls dar und weniger die Kürze, denn Primingeffekte findet man bereits bei Intervallen von 80 ms.

Welche Konsequenzen ergeben sich nun aufgrund des Problems des Holismus? Verschiedene Theoretiker geben hier unterschiedliche Antworten:

- Für Duhem besteht die zentrale Schlussfolgerung darin, dass Theorien nicht nur aufgrund ihres empirischen Erfolgs verglichen werden sollten, sondern dass andere Kriterien, wie z.B. Einfachheit, Allgemeinheit oder Verträglichkeit mit anderen etablierten Theorien, mit heranzuziehen sind.
- Für Popper ergibt sich die Forderung, die möglicherweise problematischen Zusatzannahmen unabhängig zu testen. Falls diese die Prüfung bestehen, so kommen sie nicht mehr als Ursachen für das Scheitern der Vorhersagen in Frage. So kann z.B. in Bsp.2-24 die semantische Nähe zwischen Konzepten unabhängig bestimmt werden.
- Für Quine ergibt sich als zentrale Konsequenz, dass letztendlich alle in der Ableitung der Vorhersagen involvierten Faktoren in Zweifel gezogen werden können, auch die logischen Prinzipien, welche den Schlussfolgerungen zugrunde liegen. Man wird jedoch – aus pragmatischen Gründen – zuerst jene Annahmen einer Revision unterziehen, welche die geringsten Auswirkungen auf andere Aspekte unseres Wissens haben. So wird man nicht beim Scheitern eines Experiments sofort die logischen Prinzipien, auf welche die Herleitungen der Vorhersagen beruhen, in Zweifel ziehen.

**Bsp.2-28: Überlichtgeschwindigkeit von Neutrinos**

Am 22. Sept. 2011 wurde am Forschungszentrum CERN Neutrinos mit Überlichtgeschwindigkeit gemessen. Dieses Ergebnis widerspricht der Speziellen Relativitätstheorie.

Obwohl dieses Ergebnis ein »gefundenes Fressen« für die Presse darstellte, glauben die wenigsten Physiker, dass damit Einsteins Relativitätstheorie widerlegt ist. Solange das Ergebnis nicht durch unabhängige Messungen bestätigt wurde, gilt das Ergebnis als nicht relevant.

Andererseits gibt es schon Versuche die Spezielle Relativitätstheorie durch ad hoc Annahmen zu stützen, z.B. durch die Annahme, dass die Neutrinos eine »Abkürzung« durch den Raum genommen haben.

In jedem Falle führt das Ergebnis nicht zu einer Revision der gut bestätigten Speziellen Relativitätstheorie, denn diese hätte Auswirkungen auf unser gesamtes physikalisches Weltbild.

Alle drei Positionen machen Sinn. Denn Kriterien, wie die Einfachheit und Allgemeinheit sind für die Beurteilung von Theorien und Modellen relevant (vgl. Abschnitt **xxxx**). Es gilt jedoch auch, dass in vielen Fällen (jedoch nicht in allen), die Zusatzannahmen unabhängig geprüft werden können. In diesem Falle macht es wenig Sinn, den Grund für die fehlerhaften Vorhersagen der Theorie in Zusatzannahmen zu suchen. Schliesslich ist auch die Idee von Quine korrekt, dass es strategisch sinnvoll ist, zuerst die Ursachen für das Scheitern in jenen Annahmen zu suchen, die weniger zentral sind und deren Abänderung die wenigsten Auswirkungen besitzt.

Das Problem des Holismus wird m. E. überschätzt. Dieses Problem taucht vor allem dann auf, wenn über einen Bereich noch wenig Wissen vorhanden ist. Sobald dieser jedoch gut untersucht ist, lassen sich sehr schnell verschiedene ad hoc Annahmen ausschliessen (Siehe hierzu auch die Bemerkungen am Ende des Kapitels).



Bemerkung zur Duhem-Quine-These:

Die Holismus-These wird oft auch *Duhem-Quine-These* oder als *Quine-Duhem-These* bezeichnet. Diese Bezeichnung ist m. E. irreführend, da Quine eine viel radikaler These vertrat als Duhem (die gleiche Position vertritt Lakatos, 1978; S. 97).

Gemäss der Konzeption von Quine können Theorien nicht unabhängig von unserem gesamten Wissen getestet werden. Das bedeutet, dass nur die Wissenschaft als Gesamtheit getestet werden kann. Eine derartige radikale Position liegt Duhem völlig fern.

Es sei noch ausgeführt, dass Quine diese radikale These später revidierte (vgl. hierzu Curd & Cover, 1998).

Als nächstes betrachten wir ein Problem, das sehr eng mit dem Holismus-Problem verwandt ist und ebenfalls gegen eine radikale Position der Falsifikation spricht.

2.3.1.2 DIE UNTERBESTIMMTHEIT VON THEORIEN DURCH DATEN

Zur Explikation der grundlegenden Idee führen wir ein Modell ein, den *Schwarzen Kasten (Black Box)*.

2.3.1.2.1 Die Theorie des Schwarzen Kastens

Gegeben sei ein Kasten. Aus diesem hängen eine Menge Drähte, von denen einige mit *Eingang* und andere mit *Ausgang* markiert sind (Abb. 2-8:). Der Kasten kann nicht geöffnet werden, da dies zu seiner Zerstörung führen würde.

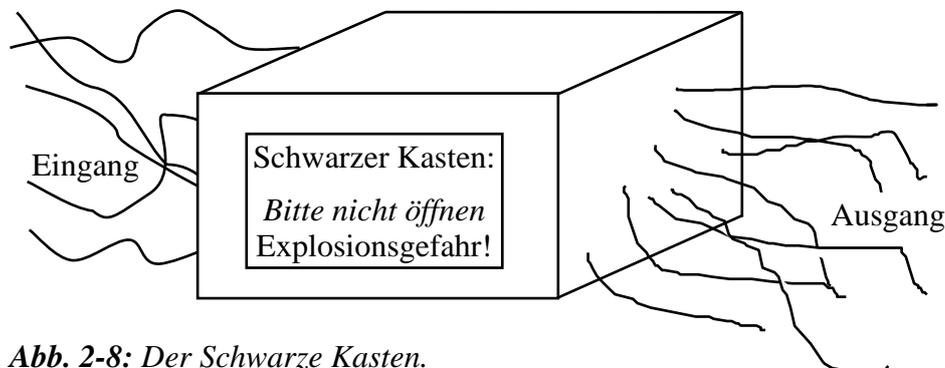


Abb. 2-8: Der Schwarze Kasten.



Bemerkung zum Begriff des Schwarzen Kastens:

Der Kasten heisst *Schwarzer Kasten*, weil er nicht geöffnet werden kann und sein daher Inhalt verborgen bleibt. Die Eigenschaft schwarz bezieht sich daher auf den Inhalt, der im Dunkeln bleibt.

Der Schwarze Kasten ist eine Metapher für die Natur.

Zwei Wissenschaftlerinnen, nenne wir sie *E* (Empiristin) und *T* (Theoretikerin) erhalten den Auftrag, möglichst viel über den Schwarzen Kasten in Erfahrung zu bringen.

Forscherin *E* stellt folgende Überlegung an:

Leider kann ich den Kasten nicht öffnen, um zu erforschen, wie er im Inneren aufgebaut ist. Zum Glück gibt es diese Ein- und Ausgänge, die es mir erlauben, zu untersuchen, wie sich das Ding verhält.

Aufgrund dieser Überlegung beginnt sie nun, an die Eingangsdrähte bestimmte Signale anzulegen und Messungen an den mit *Ausgang* markierten Drähten vorzunehmen.

Als Ergebnis erhält er eine Liste von *Input-Output – Kombinationen*, bzw. *Stimulus-Response – Paare (S-R-Paare)*:

Stimulus	Response
S_1	R_1
S_2	R_2
...	...
S_n	R_n

Nachdem *E* alle möglichen *S-R-Paare* untersucht hat, hält sie ihre Aufgabe für erfolgreich gelöst, d.h. er glaubt alles Wissenswerte über den Schwarzen Kasten herausgefunden zu haben, denn Sie kann nun für jede Kombination von Stimuli S_i vorhersagen, wie der Kasten reagieren wird, d.h. welche Kombination der R_j am Ausgang auftreten wird.

Die zweite Forscherin *T* wendet zwar – *oberflächlich betrachtet* – eine ähnliche Strategie wie *E* an, hat jedoch eine fundamental unterschiedliche Einstellung zum Schwarzen Kasten bzw. was das Ziel ihrer Forschung betrifft. Diese lässt sich wie folgt zusammenfassen:

Der Kasten besitzt einen inneren Mechanismus (siehe Abb. 2-9:). Meine Aufgabe ist es, durch gezielte Experimente diese innere Struktur aufzudecken. Die beste Vorgehensweise besteht daher darin, bestimmte Annahmen über die innere Struktur des Kastens zu formulieren. Aus diesen leite ich dann Vorhersagen ab, die ich mittels gezielter Experimente untersuche.

Das Ziel meiner Forschung muss daher darin bestehen, mittels eines möglichst einfachen und allgemeingültigen Mechanismus das bisher beobachtete Verhalten zu erklären und das zukünftige vorherzusagen.

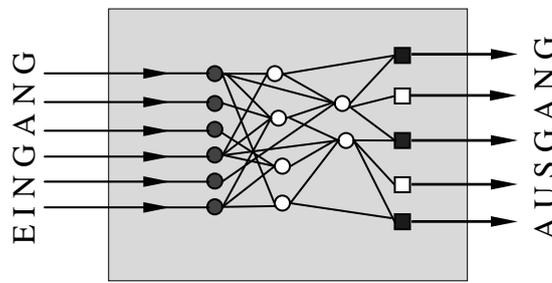


Abb. 2-9: Schwarzer Kasten mit innerem, verborgenem Mechanismus.

Man beachte, dass die Forscherinnen *scheinbar* gleich vorgehen, dass jedoch in Wirklichkeit völlig unterschiedliche Untersuchungsstrategien verwendet werden. Während *E* systematisch alle Input-Output Relationen untersucht, geht *T* selektiv vor, indem sie vor allem jene Input-Output Relationen ins Auge fasst, die zur Prüfung der Theorie relevant sind. *T* kann daher viel besser mit dem Problem der *kombinatorischen Explosion* der möglichen Testsituationen umgehen. Damit ist gemeint, dass die Anzahl möglicher Tests sehr schnell mit der Anzahl vorhandener Ein- und Ausgänge steigt.



Bemerkung zur »kombinatorischen Explosion« der Testsituationen:

Beide Forscherinnen sehen sich noch einem zusätzlichen Problem gegenüber, nämlich der kombinatorischen Explosion der Anzahl Testsituationen.

Nehmen wir an, es gäbe 100 Eingänge, mit 2 möglichen Inputs (Strom anlegen oder nicht). Dann gibt es bereits 2^{100} mögliche Testsituationen. Dies ist um einiges mehr als die Anzahl Sekunden, welche das Weltall bisher existiert.

Diese kombinatorische Explosion ist für Forscherin *E* viel problematischer als für *T*, da es erster um die genaue Tabellierung der Input-Output Paare geht, während dies für letztere von untergeordneter Bedeutung ist.

Die Vorgehensweise von T besitzt noch einen weiteren Vorteil: E kann – da sie keine Theorie besitzt – keinerlei Vorhersagen bezüglich neuer Testsituationen machen, während dies für T ein explizites Ziel darstellt.

Aufgrund der hier gegebenen Darstellung werden sich vermutlich die meisten eher mit Forscherin T als mit E identifizieren. Daher ist es wichtig, das *zentrale Problem*, das mit einer derartigen Einstellung verbunden ist, näher zu betrachten.



Prinzip 2-12: *Das Problem der Unterbestimmtheit von Theorien durch Daten:*

Zu jeder Menge von Daten lassen sich immer mindestens zwei unverträgliche Theorien finden, welche die Daten gleich gut erklären.

Im Kontext des Schwarzen Kastens besagt das Problem der Unterbestimmtheit, dass es zwei Theoretikerinnen T_1 und T_2 gibt, welche unterschiedliche Mechanismen postulieren, um die vorhandenen Daten zu erklären. Bevor wir konkrete Beispiele aus der psychologischen Forschung betrachten, wollen wir das Prinzip anhand eines ganz einfachen Beispiels demonstrieren.



Bsp.2-29: *Testung logischer Schaltkreise:*

Gegeben: Zwei Schwarze Kästen mit zwei Eingängen und einem Ausgang (Abb. 2-10).

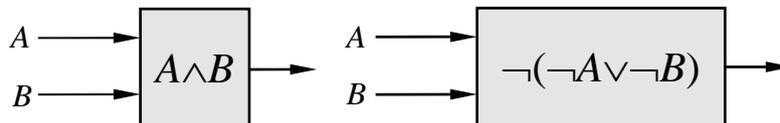


Abb. 2-10: *Zwei Schwarze Kästen mit identischem Input-Output Verhalten.*

An den Eingängen kann entweder Strom angelegt werden (= 1) oder nicht (= 0). Am Ausgang erhält man entweder eine 1 oder eine 0. Damit ergeben sich 4 mögliche Kombinationen von Eingaben: (0, 0), (0,1), (1,0), (1, 1).

Der linke Schwarze Kasten repräsentiert die logische UND-Verknüpfung. Daher wird am Ausgang nur im Falle von (1, 1) als Input eine 1 erscheinen und in allen anderen Fällen eine 0.

Nun gilt jedoch nach den Regeln von DeMorgan, dass der rechte Schwarze Kasten exakt die gleiche Antwort auf die vier Inputkombinationen liefert. Dies lässt sich auf einfache Weise mit Hilfe von Wahrheitstafeln zeigen (Übung 2-5)

Die beiden Schwarzen Kästen zeigen daher exakt das gleiche Input-Output Verhalten, obwohl die Art und Weise, wie der Input in den Output überführt wird, auf unterschiedlichen Mechanismen beruht.



Bemerkung zu den Wahrheitstafeln für aussagenlogische Ausdrücke

Die Wahrheitstafeln geben für die logischen Junktoren (Verknüpfungen) \neg , \wedge , \vee , \rightarrow , \leftrightarrow für alle möglichen Wertebelegungen der Ausgangsvariablen an, welche Wahrheitswert sich als Ergebnis der Anwendung des Junktors ergibt. Für die einzelnen Junktoren ergibt sich die folgende Wahrheitstafel (Tab. 2-2):

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftrightarrow B$
1	1	0	1	1	1	1
1	0	0	0	1	0	0
0	1	1	0	1	1	0
0	0	1	0	0	1	1

Tab. 2-2: Wahrheitstafel für die aussagenlogischen Verknüpfungen (Junktoren).

Die ersten beiden Spalten geben alle möglichen Kombinationen von Werten wieder, welche die beiden Variablen A und B annehmen können.

Spalte 3 zeigt die Wahrheitswerte, die sich bei Anwendung von » \neg « auf A ergeben (Die Wahrheitswerte werden einfach umgekehrt).

Analog zeigen die Spalten 4-7, die Wahrheitswerte der Aussagen, welche sich aus A und B durch Anwendung der Verknüpfungen » \wedge «, » \vee «, » \rightarrow «, und » \leftrightarrow « ergeben.

Mit Hilfe der Wahrheitstafeln lassen sich die Werte von komplexeren logischen Ausdrücken unter den verschiedenen Wertebelegungen der Ausgangsvariablen ermitteln, indem man die Formel aus den Teilformeln durch systematische Einführung von Junktoren aufbaut

Das folgende einfache Beispiel demonstriert das Prinzip.



Bsp.2-30: Anwendung der Wahrheitstafelmethode zur Bestimmung der Werte einer Formel unter verschiedenen Belegungen der Ausgangsvariablen

Es soll gezeigt werden, dass der Ausdruck $A \rightarrow B$ identisch mit dem Ausdrücken $\neg A \vee B$, sowie $\neg A \vee (A \wedge B)$, d.h. die drei Ausdrücke besitzen unter den verschiedenen Wertebelegungen der Ausgangsvariablen A und B exakt die gleichen Werte.

Die Spalten 1 und 2 von Tab. 2-3 enthalten alle möglichen Kombinationen von Wertebelegungen der beiden Ausgangsvariablen.

Spalte 3 enthält die Werte bei Negation von A . Da $\neg A$ einen Teil der komplexeren Formel bildet, wird für $\neg A$ eine eigene Spalte gebildet (nicht aber für $\neg B$, das nicht Teil der komplexen Formel ist).

A	B	$\neg A$	$\neg A \vee B$	$A \wedge B$	$\neg A \vee (A \wedge B)$
1	1	0	1	1	1
1	0	0	0	0	0
0	1	1	1	0	1
0	0	1	1	0	1

Tab. 2-3: Wahrheitstafel zur Bestimmung der Werte der Ausdrücke $\neg A \vee B$ und $\neg A \vee (A \wedge B)$.

Spalte 4 enthält die Werte der komplexen Formel. Diese ergeben sich durch Anwendung der logischen ODER-Verknüpfung auf die beiden Spalten $\neg A$ und B (vgl. Tab. 2-2).

Spalte 5 enthält die Werte von $A \wedge B$ unter den verschiedenen Belegungen. (vgl. Tab. 2-2).

Spalte 6 enthält die Werte des Ausdrucks $\neg A \vee (A \wedge B)$ unter den verschiedenen Belegungen von A und B . Diese Werte ergeben sich durch Anwendung der ODER-Verknüpfung auf die Einträge in Spalte 3 und 5.

Vergleicht man die Werte in der vierten und der letzten Spalte von Tab. 2-3 mit den Einträgen in Tab. 2-2 in Spalte $A \rightarrow B$, so sieht man, dass für alle drei Formeln exakt die gleichen Werte für alle Kombinationen der Belegung von A und B vorliegen. Die Ausdrücke sind daher logisch äquivalent.



Bemerkung zur Plausibilität den Wahrheitstafeln:

Die Wahrheitstafeln geben in den meisten Fällen die alltägliche Interpretation der logischen Junktoren wieder:

- $A \wedge B$ ist genau dann wahr, wenn sowohl A wie auch B wahr ist.
- $A \vee B$ ist genau dann wahr, wenn entweder A oder B wahr ist oder beide Teilaussagen wahr sind.
- $A \leftrightarrow B$ ist genau dann wahr, wenn die Wahrheitswerte der beiden Teilaussagen identisch sind.
- Eine Ausnahme bildet die logische Implikation $A \rightarrow B$. Diese Aussage ist genau dann falsch, wenn A wahr ist und B falsch. Dies entspricht nicht vollständig unserer Intuition.

Kehren wir zum Problem der Unterbestimmtheit von Theorien durch Daten zurück. In der Psychologie wurde diese Problematik ausgiebig diskutiert und in der Tat gibt es in vielen Bereichen der Psychologie alternative Interpretationen von bestehenden Daten. In der Psychologie gab es vor einigen Jahren ausgiebige Diskussionen bezüglich der Unterbestimmtheit von kognitiven Repräsentationen und Prozessen.

Die erste Diskussion betraf die Frage der eindeutigen Identifikation von Repräsentationen im Kontext der Debatte über den Status visueller Vorstellungen. Eine Gruppe von Forschern behauptete, dass visuelle Vorstellungen auf einer speziellen »bildhaften« Repräsentation beruhen (Kosslyn, Pinker, Smith, & Schwartz, 1979), während Pylyshyn (1979a, 2002) die Auffassung vertrat, dass keine der für eine bildhafte Repräsentation vorgebrachten Argumente und Befunde eine derartige spezielle Form der Repräsentation rechtfertige.

Anderson (1978) behauptete, dass es prinzipiell nicht möglich sei, mit Hilfe von Verhaltensdaten zwischen verschiedenen Formen von Repräsentationen zu unterscheiden. Er demonstrierte, dass es im Prinzip möglich ist, jede Repräsentation durch eine andere zu simulieren, so dass beide bei gleichem Input zu exakt dem gleichen Verhalten führen, ähnlich wie im Bsp.2-29 die logische Schaltung $A \wedge B$ durch die Schaltung $\neg(\neg A \vee \neg B)$ simuliert werden kann.



Bemerkung zu Repräsentationen und Prozesse:

Die gegebene Darstellung vereinfacht etwas die Diskussion, denn es ist nicht möglich, Repräsentation unabhängig von den darauf arbeitenden Prozessen zu betrachten: Die Auswirkungen einer speziellen Form der Repräsentation auf das Verhalten hängt von den Prozessen ab, welche die Repräsentation generieren, transformieren etc.

Eine zweite wichtige Diskussion betraf die Frage der seriellen vs. parallelen Verarbeitung von Information.



Bsp.2-31: *Serielle vs. parallele Verarbeitung: Das Experiment von Sternberg (1966)*

Zielsetzung:

Das Experiment untersuchte die Frage des Zugriffs auf Information im Kurzzeitgedächtnis.

Stimulusmaterial:

Das Material bestand aus den Ziffern 0-9.

Durchführung (Experiment 1):

Der Vp wurde in jedem Durchgang zuerst eine Menge von Ziffern nacheinander präsentiert. Diese Menge (= *Gedächtnismenge*) variierte in ihrer Grösse zwischen 1 und 6. Die Vpn wurden gebeten, die Ziffern im Gedächtnis zu behalten. Anschliessend wurde eine einzelne Ziffer als Testitem präsentiert und die Vp musste möglichst schnell entscheiden, ob das Testitem in der Gedächtnismenge enthalten war.

Ergebnisse:

Das Experiment lieferte zwei zentrale Ergebnisse:

1. Die Reaktionszeit steigt linear mit der Grösse der Gedächtnismenge.
2. Der Anstieg ist identisch für positive Testitems (= Testitems, die in der Gedächtnismenge enthalten waren), wie für negative Testitems (= Testitems, die in der Gedächtnismenge nicht enthalten waren).

Interpretation der Ergebnisse durch Sternberg:

1. Die Gedächtnismenge wird seriell auf das Testitem hin durchsucht.
2. Die Suche ist *exhaustiv* (und nicht *selbst-terminierend*), d.h. es wird immer die gesamte Gedächtnismenge durchsucht (und die Suche wird nicht abgebrochen, sobald eine Übereinstimmung gefunden wurde).

Bemerkung:

Die Annahme der seriellen Absuche der Gedächtnismenge scheint plausibel. Die Annahme einer vollständigen Suche scheint hingegen auf den ersten Blick ineffizient und daher wenig plausibel. Warum sollte die gesamte Menge abgesucht werden, wenn sich das Testitem in der Gedächtnismenge z.B. an der 1. Position befindet?

Sternberg argumentiert wie folgt. Das Absuchen der Gedächtnismenge verläuft sehr schnell. Würde man nun nach jedem inspizierten Item den Suchprozess unterbrechen um zu prüfen, ob das Testitem vorhanden ist, so würde dies den gesamten Suchprozess extrem verlangsamen. Daher wird zuerst die gesamte Menge durchsucht und erst dann entschieden.

Gemäss Sternberg handelt es sich bei der Suche im Kurzzeitgedächtnis (KZG) um einen *seriellen* und *exhaustiven* Prozess.

Die Ergebnisse lassen sich jedoch auch mit Hilfe eines *parallelen Zugriffsprozess* erklären: Demzufolge wird gleichzeitig auf alle Items der Gedächtnismenge zugegriffen. Dieser Zugriffsprozess verlangsamt sich jedoch in Abhängigkeit von der Anzahl Items in der Gedächtnis-

menge: Je mehr Items in der Gedächtnismenge vorhanden sind, desto schwieriger und langsamer ist der Zugriff auf ein einzelnes Item (= *Fächereffekt*).

Die Ergebnisse von Sternberg lassen sich also sowohl mittels einer exhaustiven seriellen Suche als mittels eines kapazitätsbeschränkten parallelen Suchprozesses erklären. Der Ausdruck *kapazitätsbeschränkt* bezieht sich hier auf die Tatsache, dass die Effizienz des Zugriffs von der Anzahl Items in der Gedächtnismenge abhängt.

Wir haben hier ein schönes Beispiel für alternative Erklärungen eines Phänomens. Derartige Alternativerklärungen sind in der Psychologie häufig anzutreffen. In Kapitel 3 wird das Problem von Alternativerklärungen in der Psychologie erneut aufgenommen und es werden weitere Beispiele diskutiert. Wir diskutieren als nächstes mögliche Strategien im Umgang mit dem Problem des Holismus und der Unterbestimmtheit.

2.3.1.3 STRATEGIEN IM UMGANG MIT HOLISMUS UND UNTERBESTIMMTHEIT VON THEORIEN DURCH DATEN

Grundsätzlich lassen sich drei Einstellungen gegenüber dem Problem der Unterbestimmtheit identifizieren:

1. *Vermeidung des Problems der Unterbestimmtheit:*

Diese Position hält das Problem der Unterbestimmtheit für prinzipiell unlösbar. Daher werden alle Arten von Spekulationen über den inneren Aufbau des Schwarzen Kastens vermieden (Dies entspricht der Position der Forscherin *E* in der obigen Darstellung).

2. *Nichteliminierbarkeit der Unterbestimmtheit:*

Das Problem der Unterbestimmtheit wird als grundlegende Einschränkung der möglichen Erkenntnis bezüglich des inneren Aufbaus des Schwarzen Kastens gesehen, welche *grundsätzlich nicht eliminierbar* ist. Aus diesem Grunde wird angenommen, dass die Forschung an einen Punkt gelangt, wo zwischen unterschiedlichen Theorien nicht mehr unterschieden werden kann. Ungeachtet dessen wird die Theoriebildung als ein wertvoller Bestandteil der Erforschung des Schwarzen Kastens gesehen. Ein prominenter Vertreter dieser Position ist John Anderson (Anderson, 1978, 1979, 1990).

3. *Pragmatische Lösbarkeit des Problems:*

Das Problem der Unterbestimmtheit wird als Symptom dafür betrachtet, dass die Theorien nicht ausreichend genau formuliert sind. Je exakter und umfassender eine Theorie ist, desto schwieriger wird es, eine konkurrierende zu finden, die Gleiches leistet. Daher verliert im Zuge der Erforschung und der Formulierung exakterer Theorien das Problem in der Praxis immer weiter an Bedeutung (auch wenn es als grundsätzliches Problem bestehen bleibt). Diese Position, welche von einer Reihe von kognitiven Wissenschaftlern ver-

treten wird (z.B. Massaro & Friedman, 1991; Pylyshyn, 1979b), soll am Beispiel der Schaltkreise verdeutlicht werden.



Bsp.2-32: Unterscheidung zwischen logischen Schaltkreisen
(Fortsetzung von Bsp.2-29):

Die beiden Schwarzen Kästen ($A \wedge B$) und $\neg(\neg A \vee \neg B)$ in Bsp.2-29 bleiben ununterscheidbar, solange nur die Teststrategie »Wahrheitswerte (Strom) an dem Eingängen setzen und den Wahrheitswert am Ausgang ablesen« verwendet wird. Um die beiden *Black Boxes* zu unterscheiden, müssen neue, verfeinerte Prüfverfahren gesucht werden.

Eine Möglichkeit besteht darin, Annahmen über die *Zeit* zu spezifizieren, die für die Erzeugung des Ausgangssignals benötigt wird. Man kann z.B. Annahmen darüber machen, wie viel Zeit die einzelnen elementaren Operationen \neg , \wedge und \vee benötigen, und genau festlegen, wie diese Operatoren zeitlich abgestimmt sind. Daraus liessen sich dann die Reaktionszeiten berechnen, die aufgrund der unterschiedlichen Komplexität der logischen Schaltungen unterschiedlich sein müssten.

Die Ursachen für Underdeterminiertheit und Holismus sind m. E. im ungenügenden Wissen über einen Objektbereich zu finden. Das fehlende Wissen hat zwei wichtige Konsequenzen:

1. Die entwickelten Modelle sind oft nur sehr skizzenhaft formuliert und lassen sich folglich nicht genau testen.
2. Die Theorien aus verschiedenen Bereichen des Gebietes sind nicht ausreichend verbunden, sodass sich keine wechselseitigen Restriktionen durch die verschiedenen Theorien ergeben.

Beide Aspekte sind wichtig, denn wenn die Modelle präzise Vorhersagen machen, so lassen sie sich auf einfache Weise testen und mit Alternativmodellen vergleichen. Dies wird dazu führen, dass es immer schwieriger wird, gute Alternativmodelle bzw. -theorien zu bestehenden erfolgreichen Modellen zu finden.

Ebenso hilfreich ist es, wenn Theorien aus unterschiedlichen Teilbereichen sich gegenseitig beschränken. Hilary Putnam vergleicht in einem BBC Interview (siehe YouTube) die Wissenschaft mit einem Kreuzworträtsel, wo die einzelnen Teile zusammenpassen müssen. Dies ist für reife Wissenschaften der Fall. Daher stellt sich diesen Wissenschaften im Allgemeinen weder das Problem der Unterbestimmtheit noch das Problem des Holismus: Sobald ein Gebiet theoretisch durchdrungen ist, wird es immer schwieriger alternative Theorien und Annahmen zu finden, die alle vorliegenden Daten erklären können.

Aufgrund dieser Überlegungen ist der pragmatische Weg zur Lösung der beiden Probleme m. E. der einzig richtige: Es müssen immer genauere Modelle und Theorien formuliert werden, welche möglichst

viele empirische Daten erklären. Zusätzlich sollten die Theorien aus verschiedenen Bereichen vernetzt sein, sodass sie Beschränkungen für den Entwurf neuer Theorien in benachbarten Bereichen liefern.

Die Tatsache, dass es innerhalb der Psychologie in fast jedem Teilbereich alternative Erklärungsansätze gibt, zeigt, dass diese Wissenschaft möglicherweise noch nicht den Reifegrad anderer Disziplinen (Physik, Chemie, oder Biologie) erreicht hat.



Bemerkung zum Reifegrad von Grenzgebieten der Physik

Man mag einwenden, dass sich in Grenzbereichen der Physik wie z.B. der Kosmologie oder der Theorie Schwarzer Löcher durchaus das Problem der Unterbestimmtheit zu finden ist.

Dies ist korrekt. Der Grund hierfür ist eindeutig darin zu finden, dass es in diesen Bereichen zwar viele Spekulationen aber keine oder nur wenig Daten gibt, um diese zu testen (siehe z.B. Susskind, 2010).

Betrachten wir als Nächstes zwei andere Erklärungsansätze der Funktionsweise der Wissenschaft, welcher der Falsifikation einen geringen Stellenwert beimisst.

2.3.2 Die Logik der Forschungsprogramme von Lakatos und die Idee des Paradigmenwechsels von Kuhn

In diesem Abschnitt werden zwei Ansätze behandelt, welche nicht nur die Funktion von Wissenschaft zu erklären suchen, sondern auch die Entwicklung von wissenschaftlichen Theorien.

2.3.2.1 DIE LOGIK DER FORSCHUNGSPROGRAMME NACH LAKATOS

Gemäss der Konzeption von Imre Lakatos (1922–1974) [eigentlicher Name: Imre Lipschitz] ist bei Theorien zwischen einem *Theoriekern* und *Zusatzannahmen* zu unterscheiden. Der Theoriekern enthält die zentralen theoretischen Annahmen, während die Zusatzannahmen eine Art *Schutzgürtel* um den Theoriekern bilden. Werden die Vorhersagen einer Theorie durch die Daten nicht bestätigt, so wird versucht, die Zusatzannahmen so zu verändern, dass der Theoriekern bestehen bleibt.



Bsp.2-33: Theoriekern und Zusatzannahmen der assoziativen Lerntheorien:

Der Theoriekern der assoziativen Lerntheorien besteht in der Annahme, Lernprozesse bestünden in der Bildung von Assoziationen zwischen Stimuli oder zwischen Stimuli und Responses.

Eine Zusatzannahme besteht darin, dass die Assoziationen zwischen Stimulus und Response im Lernprozess gebildet werden, indem Stimulus und Response nur oft genug gepaart, d.h. in raum-zeitlicher Nähe präsentiert werden.

Das Phänomen der Blockierung (vgl. Bsp.2-26) zeigt, dass die Theorie in der dargestellten Form zu falschen Vorhersagen führt: In diesem Fall wird ein Stimulus nicht mit einem anderen assoziiert, auch wenn beide beliebig oft gepaart werden (Kamin, 1969)

Diese Erkenntnis führte nicht zur Ablehnung der zentralen theoretischen Annahme sondern zur Entwicklung eines neuen Lernprinzips, der so genannten *Rescorla-Wagner-Regel* oder *Delta-Regel* (Rescorla & Wagner, 1972):

Ein neuer Stimulus wird mit einem anderen nur dann assoziiert, wenn seine prädiktive Potenz über die der anderen vorhandenen Stimuli hinausgeht. Dies bedeutet, dass der neue Stimulus zu einer neuen Vorhersage über den zu assoziierenden Stimulus führen muss, welche sich nicht aus den bereits vorhandenen Stimuli erschliessen lässt.

Im Falle der Blockierung einer Ursache durch eine andere ist die blockierte Ursache immer mit der anderen Ursache, deren Wirkung bereits etabliert ist, konfundiert (d.h. beide treten immer gemeinsam auf oder es sind beide abwesend). Daher enthält die blockierte Ursache keine zusätzliche Information über die Wirkung.

Theorien können – gemäss Lakatos – nicht falsifiziert werden. Nach Lakatos wird eine etablierte Theorie nicht aufgrund von falsifizierender Evidenz verworfen. Derartige Befunde werden als *Anomalien* betrachtet und man hofft, dass sich diese im Rahmen einer Weiterentwicklung der Theorie ausmerzen lassen. Das Kriterium der empirischen Falsifikation ist daher nicht der entscheidende Faktor für die Verwerfung einer Theorien.

Der wirkliche »Mörder« einer Theorie ist eine neue Theorie, welche die empirischen Befunde erfolgreich erklären kann. Solange keine neue verbesserte Theorie vorliegt, wird die alte beibehalten.

»Contrary to naive falsificationism, no experiment, experimental report, observation statement, or well-corroberated low-level falsifying hypothesis alone can lead to falsification. There is no falsification before the emergence of a better theory.« (Lakatos, 1978; S. 35).



Bemerkung zum Erklärungspotential einer neuen Theorie:

Eine neue Theorie muss nicht alle vorhandenen empirischen Befunde erklären können. Die Geschichte der Wissenschaft enthält viele Beispiele, wo eine neue Theorie wichtige Fakten nicht erklären konnte.

Beispielsweise konnte das Kopernikanische System die fehlende Sternenparallaxe (d.h. die Sterne bewegen sich – z.B. im Gegensatz zum Mond – bei Bewegung nicht mit, was sie jedoch aufgrund der Erdbewegung tun sollten) nicht erklären. Es gibt natürlich eine Sternenparallaxe. Diese ist jedoch aufgrund der ungeheuren Distanz der Sterne (der erdnaheste Stern nach der Sonne *Proxima Centauri* ist mehr als 4 Lichtjahre entfernt) sehr gering. Die Parallaxe dieses Sternes wurde erst 1838 von Bessel entdeckt.

Die zentralen Einheiten der wissenschaftlichen Entwicklung sind gemäss Lakatos nicht vereinzelte Theorien oder Hypothesen, sondern so genannte wissenschaftliche Forschungsprogramme. Nur auf dieser Ebene lässt sich die Entwicklung der (reifen) Wissenschaften adäquat verstehen.



Konzept 2-9: Wissenschaftliche Forschungsprogramme
(Lakatos, 1978)

Ein *wissenschaftliches Forschungsprogramm* umfasst 3 wesentliche Komponenten:

1. Den Theoriekern;
2. Einen Schutzgürtel von Zusatzannahmen;
3. Heuristiken bzw. methodologische Regeln zur Lösung von Problemen.

Die Heuristiken umfassen einen positiven und einen negativen Teil.

Die *negative Heuristik* besteht in der Strategie, den Theoriekern nicht durch empirische Daten zu gefährden. Dies geschieht durch Hinzufügen neuer Zusatzannahmen oder der Abänderung bestehender (siehe Bsp.2-33).

Die *positive Heuristik* besteht in der Konstruktion neuer Probleme, Modelle und Anwendungen.

In der Psychologie gibt es Ansätze in unterschiedlichen Gebieten, die als wissenschaftliche Forschungsprogramme betrachtet werden können. Im Folgenden werden vier Beispiele für mögliche Forschungsprogramme innerhalb der Psychologie dargestellt.



Bsp.2-34: Theorie der Kognitiven Dissonanz:

Der Theoriekern (Grundlegende theoretische Annahmen):
Personen besitzen verschiedene Arten von *Kognitionen*. Diese umfassen Bewertungen, Urteile, Tatsachenfeststellungen, etc.

Die Kognitionen können untereinander verträglich (*konsonant*) sein oder widersprüchlich bzw. unverträglich (*dissonant*). Das Vorliegen von dissonanten Kognitionen führt zu Spannungen, die als negativ erlebt werden. Daher versucht die Personen diese Spannungen zu reduzieren, indem die Kognitionen angepasst werden.

Zusatzannahmen:

- Zu gefühlter Dissonanz kommt es in Allgemeinen nur dann, wenn das eigene Selbst betroffen ist (z.B. wenn durch eine Handlung ein positives Selbstbild gefährdet wird).
- Die Anpassung der Kognitionen hängt von deren Wichtigkeit für die Person ab: Je unwichtiger eine Kognition, desto eher wird sie im Prozess der Dissonanzreduktion angepasst.

Anwendungen:

Ein wichtiger Aspekt der Forschung zur Kognitiven Dissonanz besteht in der Untersuchung verschiedener möglicher Anwendungen der Theorie. Hier einige Beispiele für erfolgreiche Anwendungen (für weitere Beispiele, siehe Aronson, Wilson & Akert, 2010; Kapitel 6):

- Beurteilung von Entscheidungsoptionen vor und nach einer Entscheidung: Optionen, für die man sich endgültig entschieden hat, werden positiver bewertet als vor der Entscheidung.
- Erziehung: Milde Strafen führen eher zur Einhaltung von Regeln als schwere Strafen.
- Aufwertung von Personen, denen man selbst gutes getan hat und Abwertung von Personen, denen man Böses angetan hat.
- Abhängigkeit der Bewertung eines Gutes in Abhängigkeit von der persönlichen Investition zu dessen Erlangung.
- Kulturelle Unterschiede in der wahrgenommenen Dissonanz.

Bemerkung:

Viele der untersuchten Anwendungen basieren auf dem Prinzip der *insuffizienten Begründung* eigener Handlungen. Wenn ich z.B. einer Person, die ich für wenig sympathisch halte, eine gute Tat erweise, so bin ich in Begründungsnotstand für mein Handeln. Ein möglicher Ausweg besteht in der Aufwertung der betreffenden Person. Gemäss Aronson et al. (2010) soll Benjamin Franklin (1706–1790), einer der Väter der Amerikanischen Verfassung, diese Methode bewusst eingesetzt haben, um ein Person für sich zu gewinnen.

Das Beispiel verdeutlicht einen wichtigen Aspekt von Forschung, der in dem folgenden Prinzip formuliert ist.



Prinzip 2-13: *Forschung im Kontext von Forschungsprogrammen:*

Ein wichtiger Teil der Forschung, die sich im Rahmen von Forschungsprogrammen vollzieht, besteht – neben der Prüfung der Theorie und ihrer Grenzen – in der Suche nach neuen Anwendungsbereichen der Theorie.

Dieser wichtige Aspekt von wissenschaftlicher Forschung wurde m. E. von Popper zu wenig beachtet. Für ihn besteht Forschung vorwiegend in der Prüfung von Theorien. Dies macht jedoch nur einen Bereich von Forschung aus.

Die Dissonanztheorie ist speziell für konkrete Anwendungen interessant, da sie oft Vorhersagen macht, die der Intuition zuwider laufen.



Bsp.2-35: *Aussagen eines Psychologen zur Expo-02:*

An der Schweizer Expo-02 kam es – zumindest bei einigen Ausstellungsorten – zu längeren Wartezeiten. In der Nachrichtensendung »Echo der Zeit« von Radio DRS-2 wurde daher ein Psychologe (IFAP) befragt, welche Auswirkungen die langen Wartezeiten auf die Personen hätten.

Dieser meinte, dass bei diesen langen Wartezeiten den wartenden Personen die Ausstellungen nicht mehr gefallen könne und sie daher einen schlechten Eindruck gewinnen müssten.

Die Theorie der kognitiven Dissonanz sagt genau das Gegenteil vorher, da durch die grössere Investition ein Erklärungsnotstand entsteht, der durch die Aufwertung des gewünschten Objekts beseitigt werden kann. Die Aussagen der Veranstalter bestätigten die Erklärungen der kognitiven Dissonanztheorie.

Betrachten wir ein weiteres Forschungsprogramm: das Heuristics- und Biasprogramm. Dieses erhielt vor allem durch die Arbeiten von Amos Tversky (1937–1996) und Daniel Kahneman in den 70er-Jahren einen mächtigen Auftrieb (siehe Kahneman, Slovic & Tversky, 1982; Gilovich, Griffin & Kahneman, 2002).



Bsp.2-36: *Das Heuristics- und Bias-Programm:*

Der Theoriekern (Grundlegende theoretische Annahmen):

Personen verwenden in ihren Urteilen eine Reihe von einfachen Heuristiken und Urteilsstrategien.

Zu diesen Heuristiken und Strategien gehören unter anderem:

- Die Repräsentativitätsheuristik,
- Die Verfügbarkeitsheuristik,
- Die Strategie der Verankerung und Anpassung,
- Positive Teststrategien bei der Prüfung von Hypothesen,
- Die Unterbewertung von Basisrateninformation,
- Überbewertung kleiner Stichproben,
- Ignorieren der Regression zum Mittel, etc.

Zusatzannahmen:

Die genannten Strategien führen in vielen Situationen zu verzerrten Urteilen bzw. zu Urteilsfehlern. Es gibt jedoch auch Situationen, wo viele der Heuristiken und Strategien zu zufrieden stellenden Ergebnissen bei relativ geringem Aufwand führen können.

Anwendungsbereiche:

- Wahrscheinlichkeitsurteile,
- Kausal- und Kontingenzurteile (Urteile über Zusammenhänge),
- Beurteilung der eigenen Person,
- Beurteilung anderer Personen und Gruppen,
- Gedächtnisurteile.

Ähnlich wie die Theorie der kognitiven Dissonanz hat das Heuristics- und Biasprogramm Anwendungen in den verschiedensten Bereichen. Diese betreffen nicht nur die kognitive Psychologie, sondern auch die Sozialpsychologie (siehe z.B. Nisbett & Ross, 1980), aber auch andere Bereiche, wie die Ökonomie und Politik (siehe z.B. Thaler & Sunstein, 2009).

Als nächstes Beispiel sei ein psychologisches Forschungsprogramm genannt, das in den letzten Jahren grosses Interesse erfahren hat.



Bsp.2-37: Evolutionäre Psychologie

Der Theoriekern (Grundlegende theoretische Annahmen):

Psychologische Mechanismen sind das Ergebnis von Selektionsprozessen in der evolutionären Entwicklung des Menschen. Sie wurden zur Lösung bestimmter adaptiver Probleme gebildet.

Der menschliche Geist besteht daher aus einer Menge von Modulen, die spezialisiert für die Lösung bestimmter adaptiver Probleme sind. Zum Vergleich wird oft das Schweizer Armeemesser herangezogen, welches ebenfalls aus einer Menge von Teilen besteht, die für spezielle Aufgaben geschaffen wurden.

Zusatzannahmen:

Nicht jede sinnvolle Verhaltensweise bzw. Fertigkeit ist notwendigerweise das Produkt einer adaptiven Problemlösung. Bestimmte Fertigkeiten (wie z.B. die Fähigkeit des Lesens) können als Beiprodukte betrachtet werden, die im Zuge der Lösung verwandter adaptiver Probleme mit entstanden sind.

Anwendungsbereiche:

- Partnerwahl, Eifersucht und Attraktivitätsbeurteilung von Männern und Frauen,
- Verwandtschaftsbeziehungen,
- Aggression und Genozid,
- Emotion und Emotionsausdruck,
- Entdeckung von Betrug und anderen Formen von parasitärem Verhalten,
- Geschlechterdifferenzen hinsichtlich des Vorherrschens bestimmter Merkmale und psychischer Krankheiten,
- Mögliche Ursachen für verschiedene Arten von menschlichen Verhaltensweisen und -formen, wie Kunst, Religion und wohlütigem Verhalten, etc.

Ähnlich wie bei den anderen Beispielen von Forschungsprogrammen besteht die Haupttätigkeit der Forschung in der Ausweitung der Theorie durch das Auffinden neuer Anwendungsbereiche.

Als letztes Beispiel für ein Forschungsprogramm betrachten wir einen theoretischen Ansatz, der sowohl als Theorie aber vor allem als Methode zur (theorieabhängigen) Messung mentaler Kapazitäten von grosser Bedeutung ist.



Bsp.2-38: Die Signalentdeckungstheorie

Der Theoriekern (Grundlegende theoretische Annahmen):

Die Signalentdeckungstheorie dient zur Modellierung eines Entscheiders unter Unsicherheit, bei Vorliegen von zwei oder mehr Entscheidungsalternativen. Bei den Entscheidungsalternativen handelt es sich um einander ausschliessende Sachverhalte, z.B. ob ein Stimulus vorhanden ist oder nicht, welche von zwei oder mehr Hypothese zutrifft, oder welcher Klasse ein Stimulus angehört.

Der Theorie liegen die folgenden drei Annahmen zugrunde:

1. Die Entscheidung basiert auf einer latenten Entscheidungsvariable, welche verschiedene Werte in einem kontinuierlichen Entscheidungsraum annehmen kann. Typische Entscheidungsvariablen sind:
 - Teststatistiken, wie die t - F - etc. Statistik bei Entscheidungen zwischen statistischen Hypothesen;
 - Ein Gefühl der Vertrautheit, im Falle von Entscheidungen, ob ein Objekt bereits zuvor beobachtet wurde oder nicht (= Wiedererkennung);

- Ein Gefühl der »Signalartigkeit« eines Stimulus im Falle der Entdeckung eines Stimulus innerhalb einer »verrauschten« Umgebung.
2. Die Entscheidungsvariable nimmt im Allgemeinen für jede gezogene Stichprobe bzw. in jedem (experimentellen) Durchgang einen anderen Wert an. Von Bedeutung ist daher die Verteilung der Entscheidungsvariable unter den verschiedenen Alternativen (Hypothesen, Stimuli).
Bemerkung: Falls sich die Verteilungen für die verschiedenen Alternativen nicht unterscheiden, so liefert die Stichprobe oder der aktuelle Stimulus keine Information zur Unterstützung der Entscheidung. In diesem Falle basiert letztere auf einem Prozess des *reinen Ratens*. Dieses kann jedoch durch zusätzliche Aspekte beeinflusst werden, z.B. durch unterschiedliche Belohnungen für korrekt und falsche Antworten für die verschiedenen Alternativen oder Wissen über die Auftretenshäufigkeit der einzelnen Alternativen (=Basisraten).
 3. Um zu einer Entscheidung zu gelangen, verwendet der Entscheider je nach Anzahl der Alternativen eine oder mehrere *Entscheidungskriterien*. Diese teilen den Entscheidungsraum in verschiedene Regionen, wobei jeder Region eine Entscheidungsalternative zugeordnet wird. Fällt das Entscheidungskriterium in eine bestimmte Region, so entscheidet sich die Person für die mit dieser Region verbundenen Alternative.

Zusatzannahmen:

- Die Verteilungen der Werte der latenten Entscheidungsvariable sind nicht a priori festgelegt. Sie können aus verschiedenen Familien von Verteilungen gewählt werden z.B. Normalverteilungen (siehe Abb. 2-11) oder Poissonverteilungen. Auch Mixturen aus Verteilungen sind möglich (siehe z.B. DeCarlo, 2002; Macho, 2007);
- In multidimensionalen Entscheidungsräumen können die Entscheidungsgrenzen komplexe Formen annehmen.

Anwendungsbereiche:

- Wahrnehmung;
- Gedächtnis: Wiedererkennen, Quellengedächtnis;
- Statistik: Hypothesentests;
- Klassifikation und Mustererkennung;
- Prüfung der Effektivität und Grenzen von Heuristiken;
- Erklärung und Analyse von Wahrscheinlichkeitsurteilen.

Abb. 2-11 illustriert das Modell für den Fall von 2 möglichen Alternativen bzw. 2 möglichen Stimuli. Als Verteilungen der Entscheidungsvariable unter den beiden Alternativen fungieren zwei Normalverteilungen (auch Gauss-Verteilungen genannt, daher der Name *Gauss'sches Signalentdeckungsmodell*) mit identischer Varianz, aber unterschiedlichen Mittelwerten.

Mit Hilfe des Modells lassen sich – bei Kenntnis der Lage der Normalverteilungskurven und des Entscheidungskriteriums – die Wahrscheinlichkeiten der zwei Arten von möglichen Fehlern (*Falscher Alarm* und *Falsche Zurückweisung*) und der beiden Arten von korrekten Antworten (*Treffer* und *korrekte Zurückweisung*) ermitteln. Diese Wahrscheinlichkeiten entsprechen Flächen unter den Normalverteilungskurven (vergleiche die Beschreibung zu Abb. 2-11).

Je grösser die Überlappung zwischen beiden Kurven, desto schwieriger ist es, die beiden Arten von Signalen (z.B. alte vs. neue Items) zu unterscheiden. Im Falle von Normalverteilungen mit gleicher Varianz ($\sigma_1^2 = \sigma_2^2 = 1$) wird der Grad der Überlappung einzig durch die Differenz der Mittelwerte bestimmt. Diese Differenz wird auch mit d' bezeichnet. Sie ist in diesem Falle ein direktes Maß für die Diskriminationsfähigkeit des Entscheiders: $d' = \mu_2 - \mu_1$.

Die Lage des Entscheidungskriteriums bestimmt die Tendenz des Entscheiders zu einer der beiden möglichen Antworten: Liegt das Kriterium weit links, so gibt es eine Tendenz zu den Antworten »H_A«, »Signal«, bzw. »Alt«. Befindet sich das Kriterium eher rechts, so besteht die Tendenz zu den Antworten »H₀«, »kein Signal«, bzw. »Neu«.

Mit Hilfe des Modells lassen sich daher die Diskriminationsfähigkeit und die Antworttendenz des Entscheiders messen. Diese Messung ist *theorieabhängig*, da sie die Gültigkeit des Modells voraussetzt.

Es wird angenommen, dass die Diskriminationsfähigkeit und die Antworttendenz durch unterschiedliche Faktoren beeinflusst werden: Die Diskriminationsfähigkeit wird vor allem durch die Qualität der Stimuli und der Fähigkeit des Entscheiders bestimmt. Der Entscheider hat hierauf gewöhnlich keinen Einfluss. Die Antworttendenz wird durch Faktoren wie die Auftretenshäufigkeit der verschiedenen Stimuli oder die Belohnung bei korrekter bzw. Bestrafung bei falscher Antwort beeinflusst. Es wird angenommen, dass die Antworttendenz der Kontrolle des Entscheiders unterliegt.

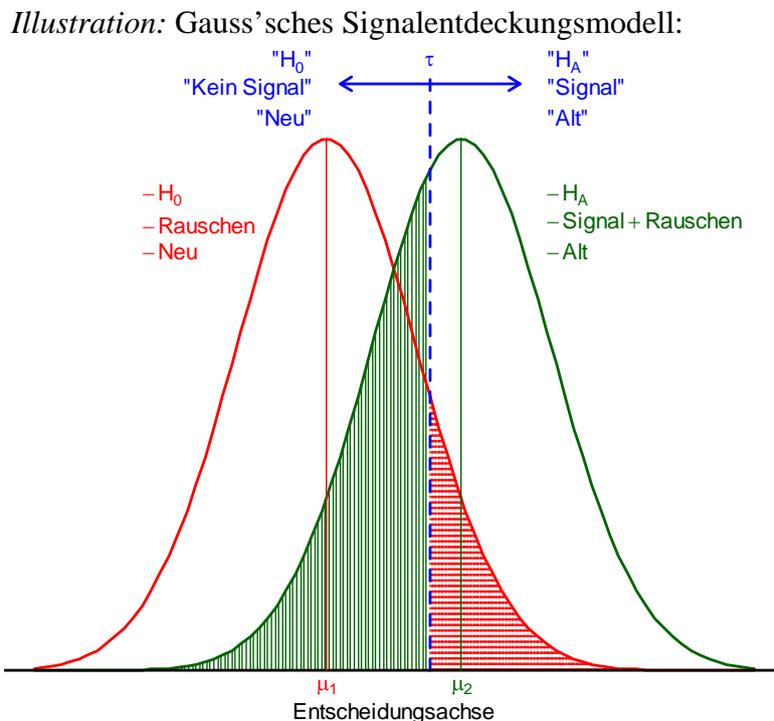


Abb. 2-11: Die grundlegenden Komponenten des Signalentdeckungsmodells: Die x -Achse repräsentiert die Entscheidungsachse. Die rote und grüne Normalverteilungskurve repräsentieren die Verteilungen der Entscheidungsvariable unter den beiden möglichen Alternativen (mit den Mittelwerten μ_1 und μ_2). Die zwei möglichen Alternativen sind H_0 vs. H_A (im Kontext des statistischen Hypothesentestens), Rauschen vs. Signal+Rauschen (im Kontext der Signalentdeckung), oder Neu vs. Alt (im Kontext des Wiedererkennens). Die gestrichelte vertikale blaue Linie, welche mit dem Symbol τ bezeichnet ist, repräsentiert das Entscheidungskriterium: Falls die Entscheidungsvariable einen Wert links des Kriteriums annimmt, so antwortet die Person mit » H_0 «, »Rauschen«, bzw. »Neu«, liegt der Wert rechts vom Kriterium, ist die Antwort » H_A «, »Signal«, bzw. »Alt«. Die rot schraffierte Fläche entspricht der Wahrscheinlichkeit eines Falschen Alarms (Typ-I Fehler oder α -Fehler). Die grün schraffierte Fläche repräsentiert die Wahrscheinlichkeit einer falschen Zurückweisung (Typ-II Fehler oder β -Fehler). Diese Wahrscheinlichkeiten werden auch mit α und β bezeichnet. Die Wahrscheinlichkeit $(1-\alpha)$ einer korrekten Zurückweisung entspricht der unschraffierten Fläche unter der roten (linken) Kurve links vom Kriterium und die Wahrscheinlichkeit $(1-\beta)$ eines Treffers entspricht der unschraffierten Fläche unter der grünen (rechten) Kurve rechts vom Kriterium.

Die dargestellten Forschungsprogramme beinhalten eine Anzahl von Gemeinsamkeiten:

1. Scheitert eine Anwendung der Theorie, so ist der Theoriekern durch diese falsifizierende Instanz nicht direkt betroffen.
2. Die Zusatzannahmen haben zwei Funktionen: (a) Sie spezifizieren Bedingungen für die Anwendungen der Theorie und (b) sie definieren einschränkende Bedingungen, die einen gewissen Schutz des Kerns bilden. Beispielsweise stellt die Zusatzannahme für die kognitive Dissonanztheorie, wonach das Selbst der Person für das Auftreten einer messbaren kognitiven Dissonanz relevant ist, eine nachträgliche Modifikation der ursprünglichen Theorie dar. Das Signalentdeckungsmodell ist ohne Zusatzannahmen, die Form der Verteilungen betreffend, überhaupt nicht prüfbar.
3. Einen wichtigen Aspekt bildet die breite Palette der möglichen bzw. erfolgreichen Anwendung der Theorien. Dies realisiert die oben erwähnte positive Heuristik, wonach ein wesentliches Ziel darin besteht, das Forschungsprogramm auf neue Probleme anzuwenden und letztere auf diese Weise zu lösen.

Gemäss Lakatos gibt es Forschungsprogramme nur für »reife« Wissenschaften. Daher nimmt er seine Beispiele aus der Physik, die er als reife Wissenschaft betrachtet. Inwieweit Lakatos die Psychologie als ausreichend reif für die Hervorbringung von Forschungsprogrammen betrachtete, ist meines Wissens nicht völlig klar. Wie aber die obigen Beispiele zeigen, enthält die Psychologie durchaus Ansätze, welche diese Bezeichnung verdienen.

Findet sich nun empirische Befunde, welche gegen die zentralen Grundannahmen des Forschungsprogramms sprechen, so führen diese *Anomalien* in der Regel nicht zur Aufgabe des Forschungsprogramms. Sie werden entweder durch Einführung weiterer Zusatzannahmen erklärt oder einfach ignoriert.



Bsp.2-39: *Befunde, welche im Widerspruch zu den Grundannahmen der Signalentdeckungstheorie stehen:*

Es gibt klare empirische Befunde, welche den Grundannahmen der Signalentdeckungstheorie widersprechen (Balakrishnan, 1999; Van Zandt, 2000). Diese Befunde wurden entweder ignoriert oder wenig überzeugend kritisiert (siehe die Diskussion zwischen Treisman (2002) und Balakrishnan & McDonald (2002)).

Ein erweitertes Modell von Mueller und Weidemann (2008) mit der neuen Zusatzannahme, dass die Entscheidungskriterien ebenfalls einer stochastischen Variation unterliegen, könnte möglicherweise die abweichenden Befunde erklären. Eine endgültige Entscheidung zu dieser Frage steht noch aus (siehe die Kritik von Balakrishnan & McDonald (2008) und die Antwort von Weidemann & Mueller (2008)).

Ein Forschungsprogramm wird so lange beibehalten, so lange es *prosperiert*. Das bedeutet, die Theorie, welche dem Programm zugrunde liegt, findet neue interessante Anwendungen bzw. sie macht neue Vorhersagen. Diese Neuanwendungen der Theorie ergeben sich meist durch Einführung neuer Zusatzannahmen.



Bsp.2-40: *Neuanwendungen des Heuristics und Bias Programms:*

Das Prosperieren des Heuristics und Biasprogramm beruht darauf, dass immer neue Heuristiken bzw. Urteilsstrategien entdeckt wurden, welche zu Urteilsverzerrungen führen können.



Bsp.2-41: *Erweiterungen des Anwendungsbereichs von Signalentdeckungsmodells:*

Die Verwendung von Mischverteilungen (d.h. Verteilungen, welche eine Mixtur von einfachen Verteilungen darstellen) führte zu erweiterten Anwendungen des Signalentdeckungsmodells: Modellierung der Quellenerinnerung (Hilford, Glanzer, Kim, & DeCarlo, 2002) oder Erklärung des Spiegeleffekts (DeCarlo, 2007). [siehe auch Bsp.2-39].



Bemerkung zur Psychologie der Forscher eines prosperierenden Forschungsprogramm :

Solange ein Forschungsprogramm prosperiert, kümmern sich die Forscher wenig um empirische Evidenz, welche der Theorie widerspricht. Vielmehr sind sie daran interessiert, die Theorie möglichst »voranzutreiben«.

»[...] the scientists attention is riveted on building his models following instructions which are laid down in the positive part of his programme. [...] If a scientist (or mathematician) has a positive heuristic, he refuses to be drawn into observations. He will 'lie down on his couch and, shut his eyes and forget about the data'. Occasionally, of course, he will ask Nature a shrewd question: he will then be encouraged by Nature's YES but not discouraged by its NO.« (Lakatos, 1978; S. 50).

Ein Forschungsprogramm *stagniert*, falls immer mehr Anomalien auftauchen, welche die Theorie nicht erklären kann. Zusätzlich gibt es immer weniger neue Anwendung bzw. Vorhersagen, welche ein Forschungsprogramm – auch im Angesicht widersprüchlicher empirischer Befunde – am Leben erhalten. Forschungsprogramme in diesem Stadium zeichnen sich vor allem durch *Rückzugsgefechte* zur Abwehr widerlegender empirischer Evidenz aus.



Bsp.2-42: *Theorie des assoziative Lernens als stagnierendes Forschungsprogramm:*

Als ein Beispiel für ein stagnierendes Forschungsprogramm mag die Theorie des assoziativen Lernens betrachtet werden. Ende der 50er- und Anfang der 60er-Jahre wurde vermehrt Evidenz gefunden, welche den Grundannahmen der assoziativen Lerntheorien widersprachen. Die Kritik betraf vor allem zwei Aspekte:

1. Es fanden sich klare empirische Befunden, welche die Annahme der graduellen Natur gelernter Assoziationen widersprechen. Dies wurde von Rock (1955) im Kontext des Paarassoziationslernens (vgl. Bsp.2-25) von Estes (1960) im Kontext der Konditionierung des Lidschlagreflexes und von Bower und Trabasso (1963) im Kontext des Kategorielernens gezeigt.
2. Die Untersuchungen von Garcia, McGowan, Ervin, und Koelling, (1968) zeigten dass räumlich-zeitliche Nähe keine notwendige Bedingung für die Bildung assoziativer Verbindungen darstellt. Deren Experimente zeigen auch, dass Assoziationen nicht zwischen beliebigen Reizen gebildet werden können. Vielmehr gibt es angeborenen Dispositionen, bestimmte Reize mit bestimmten anderen zu assoziieren: Beispielsweise wird Übelkeit eher mit dem Geruch von Futter assoziiert als mit der Form.

Diese massive Kritik führte dazu, dass die assoziativen Lerntheorien zumindest im Bereich der kognitiven Psychologie sehr schnell an Bedeutung verloren (siehe Gallistel, 1990).

In anderen Bereichen der Psychologie – wie z. B. der klinischen und der Tierpsychologie – sind assoziative Lerntheorien noch immer weitgehend akzeptiert.

Stagnierende Forschungsprogramme erhalten – analog zu den Theorien – ihren (endgültigen) »Todesstoss« durch alternative Forschungsprogramme bzw. Theorieansätze, welche die Befunde besser erklären. Im Falle der assoziativen Lerntheorien übernahm diese Rolle die kognitive Wissenschaft. Dieser zufolge besteht das Ergebnis des Lernens nicht in der Bildung von Assoziationen sondern in der Entwicklung und Verfeinerung von Repräsentationen bzw. von konzeptuellen Strukturen und Prozessen.

Es besteht jedoch die Möglichkeit, dass Forschungsprogramme, welche von einem anderen abgelöst wurde, durch neuere theoretische oder methodologische Entwicklungen wieder belebt werden (siehe hierzu die Ausführungen in Abschnitt xxx).

Offensichtlich besitzen für Lakatos empirische Befunde eine deutlich geringere Bedeutung als Motor des wissenschaftlichen Fortschritts als

für Popper, da sie weder zur Falsifikation von Theorien noch zur unmittelbaren Verwerfung von Forschungsprogrammen tauglich sind. Lakatos schätzt auch die Bedeutung von kritischen Experimenten (Experimentum crucis [siehe Abschnitt 2.2.3.1]) als viel geringer ein. Es stellt sich daher die folgende Frage:



Frage:

Welchen Stellenwert besitzen empirische Befunde im Rahmen von Lakatos Forschungsprogrammen?

Zur Beantwortung dieser Frage ist es sinnvoll, auf eine grundlegende Position von Lakatos hinzuweisen:

The case studies of this section show that rationality works much slower than most people tend to think, and, even then, fallibly. Minerva's owl flies at dusk (Lakatos, 1978, S. 87).

Gemäss dieser Auffassung führen problematische empirische Befunde nicht unmittelbar zur Aufgabe von Forschungsprogrammen. Vielmehr besteht die Bedeutung empirischer Befunde in der langsamen aber stetigen Anhäufung von Anomalien, mit denen sich ein Forschungsprogramm konfrontiert sieht. Falls diese Anomalien innerhalb eines Forschungsprogramms nicht gelöst werden können und letzteres auch keine neuen progressiven Anwendungen hervorbringt, so wird dies zur Entwicklung neuer Programme führen, welche die Probleme zumindest teilweise lösen und neue Vorhersagen ermöglichen.

Empirische Befunde haben daher im Rahmen von Forschungsprogrammen zwei wesentliche Funktionen:

1. Sie zeigen Anomalien auf, mit denen sich die gegenwärtige Theorie konfrontiert sieht.
2. Sie führen bei Anhäufung der Anomalien zur Entwicklung neuer Forschungsprogramme.

Was den Stellenwert von kritischen Experimenten betrifft, so wird – gemäss Lakatos – deren Bedeutung erst nachträglich erkannt bzw. von der Forschergemeinschaft als Experimentum crucis eingeschätzt. Als ein Beispiel nennt Lakatos das Michelson-Morley Experiment von 1881 (welches als das berühmteste gescheiterte Experiment in der Geschichte der Physik gilt). Das Experiment wurde später als Widerlegung der Theorie des Äthers betrachtet. Aber noch 1905 betrachtete Kelvin das Ergebnis nur als »[a] cloud in the clear sky«. Ebenso wenig war das Ergebnis von Michelson selbst als Widerlegung der Existenz des Äthers betrachtet worden.



Bsp.2-43: *Das Experiment von Rock (1955) versus die Befunde von Kamin (1969):*

Betrachten wir nochmals die oben präsentierten Befunde von Rock (1955) [Bsp.2-25] und Kamin (1969) [Bsp.2-33].

Beide Untersuchungen widerlegen bestimmte Aspekte der assoziativen Lerntheorie. Dennoch wurde das Experiment (bzw. die Experimente) von Rock als kritische Experimente betrachtet, während der Untersuchung von Kamin dieser Status nicht zugesprochen wurde.

Zur Erinnerung:

- Die Befunde von Rock (1955) sprechen gegen die Annahme, dass Assoziationen beim Paarassoziationslernen graduell gebildet werden.
- Die Befunde von Kamin (1969) sprechen gegen die Idee, dass räumliche und zeitliche Nähe eine hinreichende Bedingung für die Bildung von Assoziationen zwischen den zu assoziierenden Stimuli darstellt.

Ein Grund für die unterschiedliche Bewertung der beiden Experimente dürfte darin bestehen, dass die Befunde von Kamin im Rahmen des Forschungsprogramms erklärt werden konnte (durch die Entwicklung einer neuen Lernregel), während die von Rock gefundene »Anomalie« nicht zufrieden stellend erklärt werden konnte.

Ohne Entwicklung der Rescorla-Wagner Lernregel wären auch die Befunde von Kamin als kritisch betrachtet worden.

Analog verhält es sich mit den Befunden von Waldmann und Holyoak (1992). Meines Wissens gibt es bis heute keine zufrieden stellende Erklärung dieser Ergebnisse im Rahmen der assoziativen Lerntheorien (was die Befürworter assoziativer Lerntheorien dazu veranlasst, diese Befunde einfach zu ignorieren).

Nach dieser ausführlichen Darstellung der Idee der Forschungsprogramme von Lakatos wenden wir uns einer weiteren einflussreichen Konzeption von Wissenschaft zu.

2.3.2.2 DIE IDEE DES PARADIGMENWECHSELS NACH KUHN

Die Bedeutung von Thomas Kuhn (1922–1996) auf die Wissenschaftstheorie des 20. Jh. kann kaum überschätzt werden. Sein im Jahre 1962 erstmals publiziertes Hauptwerk *Die Struktur wissenschaftlicher Revolutionen* wurde mehr als 1 Million Mal verkauft. Dieses Werk fand vor allem bei Geistes- und Sozialwissenschaftlern grossen Anklang, obwohl Kuhn selbst ausgebildeter Physiker war.

Im Jahre 1965 wurde von Lakatos an der London School of Economics (wo Popper und Lakatos lehrten) ein Treffen von Kuhn und Anhängern mit Popper und seinen Anhängern organisiert (Lakatos & Musgrave, 1974). Lakatos Idee der Forschungsprogramme wurde von Kuhn inspiriert und kann als eine Art rationale Interpretation oder Rekonstruktion von Kuhns Auffassung betrachtet werden, welche die Konzeptionen von Popper und Kuhn einander annähert.

Das zentrale Konzept des Ansatzes von Kuhn ist das Paradigma.



Konzept 2-10: Paradigma:

Ein *wissenschaftliches Paradigma* umfasst die grundlegenden Annahmen, Methoden, Probleme und Lösungsansätze, die von einer Forschergemeinschaft zu einem bestimmten Zeitpunkt akzeptiert werden.

Paradigmen müssen als umfassendere Einheiten betrachtet werden als Lakatos Forschungsprogramme. Sie umfassen nicht nur Theorien, und Methoden, sondern auch z.B. metaphysische Ansichten der Forschergemeinschaft zu Ontologie oder zu Fragen der Methodologie.



Bsp.2-44: Der Behaviorismus als Paradigma:

Als Beispiel für ein Paradigma in der Psychologie mag der Behaviorismus betrachtet werden, welche eine dominierende Strömung innerhalb der Amerikanischen Psychologie in der ersten Hälfte des 20. Jh. darstellte.

Der Behaviorismus zeichnet sich durch seine antitheoretische Einstellung aus, die sich in der Ablehnung aller Arten von theoretischen Konstrukten im Speziellen von mentalen Entitäten ausdrückte.

Für Kuhn gibt es zwei zentrale Phasen in der Wissenschaftsentwicklung:

1. *Phase der normalwissenschaftliche Tätigkeit:* Hier vollzieht sich die wissenschaftliche Entwicklung innerhalb eines anerkannten Paradigmas (vgl. Konzept 2-10). Diese Phase besteht vor allem in Verfeinerungen oder Verallgemeinerungen bestehender Theorien und Konzepte.
2. *Phase des Paradimenwechsel:* Diese besteht in der Ersetzung eines bestehenden Paradigmas durch ein neues, nachdem das alte Paradigma in eine Krise geraten war. Paradigmenwechsel zeichnen sich dadurch aus, dass die grundlegenden Konzepte und Methoden in Frage gestellt und durch neue ersetzt werden. In vielen Fällen treten auch völlig neue empirische Phänomene in den Vordergrund, die bisher unbeachtet blieben bzw. als irrelevant betrachtet wurden.



Bsp.2-45: Paradigmenwechsel in der Psychologie:

Innerhalb der Psychologie kann die so genannte *kognitive Wende* Ende der 50- bzw. am Anfang der 60-er Jahre als ein Paradigmenwechsel betrachtet werden.

Der Wechsel bestand in der Ablösung der *Behaviorismus* durch die *kognitive Wissenschaft* bzw. durch die *kognitive Psychologie*.

Hierbei änderte sich die *Ontologie* der Disziplin in grundlegender Weise, indem mentale Konstrukte, wie Ziele, Motive, Strategien, Heuristiken, Pläne, etc. als Teil der wissenschaftlichen Begriffssprache akzeptiert und als zentral für die Erklärung von menschlichen Verhalten betrachtet wurden.

Weiters wurden neue Forschungsmethoden entwickelt, wie z.B. die Methode des lauten Denkens oder Computersimulationen von Denk- und Handlungsprozessen.

Schliesslich wurden neue Bereiche erforscht, welche vorher als irrelevant betrachtet wurden, wie z.B. mentale Vorstellungen.



Bemerkung zum Begriff der kognitiven Wissenschaft:

Unter dem Begriff der kognitiven Wissenschaft werden verschiedene Disziplinen zusammengefasst, wie:

- Kognitive Psychologie,
- Kognitive Neuropsychologie,
- Linguistik,
- Informatik, im Speziellen Künstliche Intelligenzforschung,
- Philosophie des Geistes.

Das Verbindende dieser Disziplinen besteht in der Auffassung, dass das menschliche (und tierische) Gehirn nur adäquat verstanden werden kann, wenn er als informationsverarbeitendes System betrachtet wird.

Die folgenden Ideen Kuhns wurden heftig kritisiert:

1. *Die Theorien verschiedener Paradigmen sind inkommensurabel:*

Die innerhalb verschiedener Paradigmen entwickelten Theorien sind nicht ineinander übersetzbar, so dass keine der rivalisierenden Theorie als eindeutig überlegen betrachtet werden kann.

2. *Paradigmenwechsel haben keine rationale Grundlage:*

Paradigmenwechsel vollziehen sich indem die Theorie mit den genaueren und umfassenderen Vorhersagen gewinnt, sondern eher durch Überredung oder Suggestion bzw. – letztendlich – durch den Wechsel zu einer neuen Forschergeneration, indem die Anhänger des alten Paradigmas einfach aussterben.

3. *Der wissenschaftliche Fortschritt ist nicht kumulativ:*

Die Entwicklung der Wissenschaften kann nicht als ein Fortschritt in Richtung grösserer Wahrheitsnähe betrachtet werden, indem Sinne, dass unsere Theorien immer besser mit der Realität übereinstimmen.

Ich halte die ersten beiden Auffassungen Kuhns für nicht haltbar. Die Inkommensurabilität, auf die vor allem Feyerabend pocht, wird von den meisten Wissenschaftstheoretikern bezweifelt (siehe z.B. Watkins, 1974). Selbst wenn es die von Kuhn postulierte Inkommensurabilität

wirklich gibt, so wird ein Paradigmenwechsel nicht wesentlich durch Propaganda und Indoktrination vollzogen, sondern basiert auf Kritik an den alten Theorien, die von den Vertretern des alten Paradigmas nicht jedoch innerhalb des neuen Paradigmas adäquat beantwortet werden kann.



Bsp.2-46: *Die kognitive Wende: Chomskys Kritik an Skinner:*

Die kognitive Wende wurde wesentlich durch eine Kritik von N. Chomsky an B. F. Skinners Buch *Verbal Behavior* (mit) ausgelöst, in der klar die Grenzen des von Skinner postulierten Mechanismus des Erwerbs bzw. des Verstehens von Sprache aufgezeigt wurde. Die Kritik Chomskys lässt sich wie folgt zusammenfassen:

1. Gemäss der Konzeption von Skinner, kann Sprache durch einen endlichen Automaten (oder eine reguläre Grammatik) erklärt werden.
2. Bestimmte Strukturen der menschlichen Sprache können mit Hilfe dieses Automaten (oder dieser Grammatik) nicht erzeugt werden.

Die Kritik Chomskys an Skinners Sprachtheorie legt eine fundamentale Schwäche behavioristischer Erklärungen der menschlichen Sprachfertigkeit frei, die von der Forschungsgemeinde verstanden wurde und nicht adäquat beantwortet werden konnte.



Bemerkung zur Kritik neuronaler Modelle kognitiver Fertigkeiten:

Eine ähnliche Argumentation wie jene von Chomsky wurde in neuerer Zeit von Gallistel und King (2009) vorgebracht. Sie kritisieren vor allem aktuelle neuronale Modelle menschlicher und tierischer kognitiver Fertigkeiten.

Ihre Hauptkritik besteht darin, dass diese Modelle nicht in der Lage sind, die kombinatorischen Komplexität, denen sich ein Individuum gegenüber sieht, zu bewältigen.

Die Frage des kumulativen Charakters des wissenschaftlichen Fortschritts ist schwieriger zu entscheiden. Denn einerseits ist es unbestreitbar, dass die wissenschaftlichen Theorien zu immer genaueren und umfassenderen Vorhersagen führen. Andererseits impliziert dies jedoch nicht, dass der theoretische Begriffsapparat der Wissenschaften die Realitäten korrekt repräsentiert. Die Tatsache, dass eine Theorie zu besseren Vorhersagen führt, impliziert nicht notwendig, dass die dieser Theorie zugrunde liegende Ontologie – d.h. die von der Theorie postulierten Entitäten – der Realität besser entspricht. Zwar ist eine korrekte und überraschende Vorhersage einer Theorie ein Hinweis auf deren Korrektheit. Dies bedeutet jedoch nicht, dass die Theorie die Realität korrekt abbildet, denn auch eine falsche Theorie kann zu korrekten

Vorhersagen führen – wie unzählige Beispiel aus der Geschichte der Wissenschaft demonstrieren.

Zusammenfassend lässt sich sagen, dass die von Kuhn vertretene Position der irrationalen Entwicklung der Wissenschaften in dieser Form nicht haltbar ist. Eine adäquatere Beschreibung gibt meines Erachtens Lakatos mit seiner Konzeption der Forschungsprogramme und deren Entwicklung. Leider fand diese Idee – vermutlich, weil sie weniger spektakulär war – eine weitaus geringere Verbreitung als jene von Kuhn.

Wir betrachten als nächstes eine letzte Konzeption von Wissenschaft, die strukturalistische Konzeption. Diese hat eine gewisse Nähe zu Lakatos Konzeption.

2.3.3 Die strukturalistische Konzeption von Wissenschaft

Die strukturalistische Konzeption wurde von verschiedenen Wissenschaftstheoretikern entwickelt. Die entscheidende Vorarbeit leistete Patrick Suppes (geb. 1922), der Formalisierung von Theorien als mengentheoretische Strukturen vorschlug. Als der eigentliche Begründer gilt Joseph Sneed (geb. 1938). Im deutschen Raum wurde die Idee von Wolfgang Stegmüller (1923–1991) und zusammen mit Kollegen weiterentwickelt.

Auch innerhalb der Psychologie fand dieser Ansatz einige Aufmerksamkeit (Westmeyer; 1989, 1992). Im Folgenden werden die Grundideen, welche diesem Ansatz zugrunde liegen, vorgestellt. Eine einfach zu lesende Darstellung mit einem ausführlichen Beispiel aus dem Bereich der Psychologie findet man in Kapitel 11 von Westermann (2000).

Die zentralen Einheiten bilden gemäss der strukturalistischen Konzeption die so genannten Theorieelemente.



Konzept 2-11: Theorieelemente:

Ein *Theorieelement* ist ein geordnetes Paar: $T = \langle K, I \rangle$. Hierbei ist K der Theoriekern und I bilden die intendierten Anwendungen der Theorie.

Der Theoriekern umfasst die Repräsentation aller Entitäten, welche in die Spezifikation der Theorie eingehen. Diese Entitäten lassen sich in zwei Klassen einteilen:

1. *T-theoretische Elemente*: Hierbei handelt es sich um Elemente, die nur innerhalb der Theorie eine Bedeutung besitzen. Ihre Messung setzt die Korrektheit der Theorie voraus.
2. *T-nicht-theoretische Elemente*: Hierbei handelt es sich um Elemente, die auch ausserhalb der Theorie eine Bedeutung aufweisen.

Zusätzlich zur Repräsentation der Entitäten enthält der Theoriekern die Spezifikation der zentralen Relationen bzw. Gesetze zwischen den Elementen, welche von der Theorie postuliert werden.

Aufgrund dieser Einteilung ergeben sich drei Arten von Strukturen:

M_{pp} : *Die Menge der partiellen potentiellen Modelle:*

Diese Struktur enthält die Spezifikation aller *T-nicht-theoretischen* Entitäten der Theorie.

M_p : *Die Menge der potentiellen Modelle:*

Diese Struktur enthält zusätzlich zur Spezifikation aller T-nicht-theoretischen Entitäten auch die Spezifikation aller *T-theoretischen Elemente* der Theorie.

M : *Die Menge der Modelle:*

Diese Struktur enthält zusätzlich zur Spezifikation aller T-theoretischen und T-nicht-theoretischen Entitäten auch die Spezifikation aller durch die Theorie spezifizierten *Relationen* bzw. *Gesetzen* der Theorie.

Zusätzlich zu den verschiedenen Modellen enthält der Theoriekern noch zwei Arten von Strukturen:

GB : *Globale Beschränkungen:*

Hierbei handelt es sich um Beschränkungen von Elementen eines potentiellen Modells oder um Beschränkungen zwischen potentiellen Modellen aus verschiedenen Anwendungen.

GV : *Globale Verbindungen:*

Hierbei handelt es sich um Verbindungen zu potentiellen Modellen aus anderen Theorien.

In der folgenden Darstellung werden wir uns vor allem auf die drei Arten von Modellen und ihrer Bedeutung konzentrieren. In Abschnitt 2.3.3.4 werden wir auch kurz globale Beschränkungen und Verbindungen behandeln.

Wir beginnen mit einem Beispiel, welches die Unterscheidung zwischen den drei Typen von Modellen illustriert.



Bsp.2-47: *Potentiell partielle Modelle, partielle Modelle und Modelle des Theoriekerns des Gauss'schen Signalentdeckungsmodells:*

Wir betrachten das Gauss'sche Signalentdeckungsmodell (vgl. Abb. 2-11) in einer konkreten Anwendung.

Gegeben:

Ein Wiedererkennungsexperiment, in welchem eine V_p alte Items – d.h. zuvor gelernte Items – von neuen Items unterscheiden muss. Bei den Items kann es sich um Worte, Gesichter etc. handeln.

Es gibt daher zwei Arten von Stimuli: N = Neue Items und A = Alte Items, und die V_p hat zwei mögliche Antworten: »N« = »Neu« und »A« = »Alt«.

Das *potentielle partielle Modell* M_{pp} , welches alle relevanten *G-SDT-nicht-theoretischen Elemente* enthält (*G-SDT* steht für Gauss'sche Signalentdeckungstheorie), lässt sich wie folgt spezifizieren:

$M_{pp} = \langle S, R, T, P \rangle$, wobei gilt:

- $S = \{N, A\}$ ist die Menge der Stimuli;
- $R = \{\gg N \ll, \gg A \ll\}$ ist die Menge der Responses;
- $T = S \times R = \{(N, \gg N \ll), (N, \gg A \ll), (A, \gg N \ll), (A, \gg A \ll)\}$ ist die Menge der Kombinationen aus Stimuli und Responses. Inhaltlich lässt sich diese Menge auch wie folgt bezeichnen:
 - $(N, \gg N \ll)$ = Korrekte Zurückweisungen.
 - $(N, \gg A \ll)$ = Falscher Alarm (Fehler vom Typ I, α -Fehler).
 - $(A, \gg N \ll)$ = Falsche Zurückweisung (Fehler vom Typ II, β -Fehler).
 - $(A, \gg A \ll)$ = Treffer.
- $P = \{P(\gg N \ll | N), P(\gg A \ll | N), P(\gg N \ll | A), P(\gg A \ll | A)\}$ ist die Menge bedingter Wahrscheinlichkeiten der Responses für die verschiedenen Arten von Stimuli. Es gilt:
 - $P(\gg A \ll | N) = 1 - P(\gg N \ll | N)$ und $P(\gg A \ll | A) = 1 - P(\gg N \ll | A)$.

Das *partielle Modell* M_p enthält alle Elemente von M_{pp} und zusätzlich die *G-SDT-theoretischen Elemente* (d.h. jene Konstrukte, welche nur innerhalb des Modells eine Bedeutung haben und nur mit Hilfe des Modells gemessen werden können):

$M_p = \langle S, R, T, P, \Theta, \delta(\theta; \tau), \Phi_N(c; \mu_N, \sigma_N^2), \Phi_A(c; \mu_A, \sigma_A^2) \rangle$ mit:

- $\Theta : S \rightarrow \mathfrak{R}$ symbolisiert die *latente Entscheidungsvariable* (im aktuellen Falle handelt es sich hierbei um die Vertrautheit des Stimulus). Diese wird als Funktion betrachtet, welche jedem Stimulus einen reellen Wert (die Vertrautheit des Stimulus) zuordnet. \mathfrak{R} bezeichnet die Menge der reellen Zahlen.



Bemerkung zur Notation:

Eine Funktion f wird üblicherweise mit $f : X \rightarrow Y$ bezeichnet, wobei X den Definitions- und Y den Wertebereich bezeichnet. Die Funktion f ordnet jedem Element $x \in X$ *genau ein* Element $f(x) \in Y$ zu, wobei verschiedenen Werten von X auch das gleiche Element von Y zugeordnet werden kann.

- $\delta(\theta; \tau): \mathfrak{R} \rightarrow R$ ist eine *Entscheidungsfunktion*, welche jedem Wert θ der latenten Entscheidungsvariable eine Response aus $R = \{\text{»N«}, \text{»A«}\}$ zuordnet.

Die Entscheidungsfunktion ist wie folgt definiert:

$$\delta(\theta; \tau) = \begin{cases} \text{»N«} & \Leftrightarrow \theta \leq \tau \\ \text{»A«} & \Leftrightarrow \theta > \tau \end{cases} \quad (2-7)$$

Die Entscheidungsfunktion ordnet einem Wert der latenten Entscheidungsvariable den Wert »N« zu, falls diese kleiner gleich dem Schwellenwert τ ist, andernfalls liefert die Funktion den Wert »A«.



Bemerkung zur Notation:

Die beiden Symbole innerhalb der Klammern von $\delta(\theta; \tau)$ sind durch einen Strichpunkt getrennt. Die Symbole vor dem Strichpunkt repräsentieren Variablen, während jene nach dem Strichpunkt so genannte Parameter (siehe Konzept 2-12 und Konzept 2-13) symbolisieren.

Man beachte auch den Unterschied zwischen der latenten Entscheidungsvariable, welche mit Θ (gross Theta) bezeichnet wird, während konkrete Werte mit θ (klein theta) bezeichnet werden.

- $\Phi_N(c; \mu_N, \sigma_N^2)$ symbolisiert die kumulative Normalverteilung der Entscheidungsvariable für die neuen Items. Diese ist durch den Mittelwertparameter μ_N und den Varianzparameter σ_N^2 charakterisiert.
- $\Phi_A(c; \mu_A, \sigma_A^2)$ symbolisiert die kumulative Normalverteilung der Entscheidungsvariable für die neuen Items. Diese ist durch den Mittelwertparameter μ_A und den Varianzparameter σ_A^2 charakterisiert.



Bemerkung zur Notation:

In $\Phi_A(c; \mu_A, \sigma_A^2)$ repräsentiert c den Wert, bis zu dem die Fläche unter der Normalverteilungskurve ermittelt wird. Es gilt daher:

$$P(x \leq c | S = A) = \Phi_A(c; \mu_A, \sigma_A^2). \quad (2-8)$$

In Worten: »Die Wahrscheinlichkeit, dass ein Wert kleiner oder gleich c beobachtet wird (gegeben, dass Stimulus A präsentiert wurde), entspricht der Fläche unter der Normalverteilungskurve (für Stimulus A) bis zum Wert c .«

So entspricht z.B. in Abb. 2-11 der Ausdruck:

$$P(x \leq \tau | S = A) = \Phi_A(\tau; \mu_A, \sigma_A^2)$$

der grün schraffierten Fläche, wobei $\Phi_A(\tau; \mu_A, \sigma_A^2)$ die rechte (grüne) Normalverteilungskurve repräsentiert und τ die Lage des Entscheidungskriteriums (die blaue gestrichelte vertikale Linie).

Das *Modell M* enthält alle Elemente von M_p und zusätzlich die Spezifikation der vom Modell vorhergesagten bedingten Wahrscheinlichkeiten der verschiedenen Antworten für die einzelnen Stimuli.

$$\begin{aligned} \pi(\gg N \ll N) &= \Phi_N(\tau; \mu_N, \sigma_N^2) \\ \pi(\gg A \ll N) &= 1 - \Phi_N(\tau; \mu_N, \sigma_N^2) \\ \pi(\gg N \ll A) &= \Phi_A(\tau; \mu_A, \sigma_A^2) \\ \pi(\gg A \ll A) &= 1 - \Phi_A(\tau; \mu_A, \sigma_A^2) \end{aligned} \quad (2-9)$$

Die zentralen prüfbareren Vorhersagen der Theorie bzw. des Modells betreffen diese Antwortwahrscheinlichkeiten.



Konzept 2-12: *Variable, Parameter und Konstante:*

Ein (*Funktions-*) *Parameter* ist eine Grösse, die hinsichtlich ihrer Variabilität zwischen einer *Variablen* und einer *Konstanten* angesiedelt ist, d.h. ein Parameter ändert sich weniger oft als eine Variable. Er ist jedoch auch keine Konstante, deren Wert sich niemals ändert.

Innerhalb der Statistik spielen Parameter eine zentrale Rolle.



Konzept 2-13: *Populationsparameter und Stichprobenkennwert (Statistik):*

Ein *Populationsparameter* ist ein Kennwert, der zur Charakterisierung von Verteilungseigenschaften innerhalb einer Population verwendet wird.

Ein *Stichprobenkennwert (Statistik)* ist eine Funktion der Stichprobe.

Bsp.2-47 präsentiert – aus pädagogischen Gründen – den Theoriekern für eine konkrete Anwendung der Signalentdeckungstheorie. Die Theorie lässt sich allgemeiner formulieren, so dass die gegebene Formulierung einen Spezialfall der allgemeineren Formulierung darstellt. Im Folgenden wird die Bedeutung der einzelnen Komponenten genauer erklärt und teilweise wird gezeigt, wie die generellere Formulierung des Theoriekerns aussieht. Wir beginnen mit der Erklärung der Menge M_{pp} der partiellen potentiellen Modelle.

2.3.3.1 DIE MENGE DER PARTIELLEN POTENTIELLEN MODELLE EINER THEORIE

Die zentrale Funktion der partiellen potentiellen Modelle M_{pp} besteht in der Spezifikation der Menge aller möglichen Anwendungen des betreffenden Theorieelements.

Es scheint sinnvoll, zuerst die folgende Frage zu klären.



Frage:

Warum spricht man von Modellen und was bedeutet der Begriff »Modelle« im aktuellen Kontext?

Der Modellbegriff, wie er hier verwendet wird, ist im Kontext der modernen Logik entstanden. Ein Modell bedeutet hier nichts anderes als eine konkrete Instanz eines formalen Systems. Das folgende einfache Beispiel aus der Aussagenlogik illustriert das Prinzip.



Bsp.2-48: Modelle in der Aussagenlogik:

Gegeben: Eine syntaktisch korrekte Zeichenkette der Aussagenlogik, z.B.: $((\neg A \vee B) \leftrightarrow (A \rightarrow B))$.

Der Ausdruck »syntaktisch korrekte Zeichenkette« besagt nichts anderes, als dass die Zeichenkette nach vorgegebenen *syntaktischen Regeln* gebildet wurde. Mit Hilfe derartiger Regeln lassen sich dann unendlich viele Zeichenketten bilden.

Die syntaktischen Regeln für die Aussagenlogik lassen sich wie folgt formulieren:

1. Die Buchstaben A, B, C, \dots bilden syntaktisch korrekte Ausdrücke. Wir nennen sie die *Grundausdrücke*.
2. Ist X ein beliebiger syntaktisch korrekter Ausdruck, so ist $\neg X$ ein syntaktisch korrekter Ausdruck.
3. Sind X und Y zwei beliebig syntaktisch korrekte Ausdrücke, so sind $(X \wedge Y)$, $(X \vee Y)$, $(X \rightarrow Y)$ und $(X \leftrightarrow Y)$ syntaktische korrekte Ausdrücke.

Durch sukzessive Anwendung dieser Regeln lassen sich alle syntaktisch korrekten Ausdrücke der Aussagenlogik formulieren. Für das obige Beispiel gilt:

- A, B sind syntaktisch korrekt.
- Anwendung von Regel 2 ergibt: $\neg A$ ist syntaktisch korrekt (Die Variable X hat hier den Wert A).
- Anwendung von Regel 3 ergibt: $(\neg A \vee B)$ ist syntaktisch korrekt (Die Variable X hat hier den Wert $\neg A$ und die Variable Y den Wert B).
- Anwendung von Regel 3 ergibt: $(A \rightarrow B)$ ist syntaktisch korrekt (Die Variable X hat hier den Wert A und die Variable Y den Wert B).

- Anwendung von Regel 3 ergibt: $((\neg A \vee B) \leftrightarrow (A \rightarrow B))$ ist syntaktisch korrekt (Die Variable X hat hier den Wert $(\neg A \vee B)$ und die Variable Y den Wert $(A \rightarrow B)$).

Die so generierten Zeichenketten sind rein syntaktische Gebilde die keine Bedeutung haben. Damit sie Bedeutung erhalten müssen sie *interpretiert* werden.

Die Bedeutungen von (syntaktisch korrekt gebildeten) Zeichenketten sind die Wahrheitswerte W oder F .

Die Zuordnung von Wahrheitswerten erfolgt nun in ähnlicher Weise, wie die Bildung der syntaktisch korrekten Ausdrücke. Zuerst werden die Grundausrücke interpretiert, d.h. es werden ihnen Wahrheitswerte zugeordnet. Man spricht auch von einer Belegung der Grundausrücke.

Die Belegung der Grundausrücke nennt man ein *Modell*.

Im aktuellen Beispiel gibt es vier Modelle oder Belegungen der Grundausrücke: $\langle W, W \rangle$, $\langle W, F \rangle$, $\langle F, W \rangle$ und $\langle F, F \rangle$.

Zur Ermittlung der Wahrheitswerte für komplexere Ausdrücke verwendete man die Wahrheitstafeln (Bsp.2-30). Hierbei handelt es sich um so genannte *semantische Regeln*, die analog zu den syntaktischen Regeln die Zuordnung von Wahrheitswerten für komplexe Ausdrücke bestimmen. Die semantischen Regeln für die Aussagenlogik lauten (wir nehmen an, dass alle Ausdrücke syntaktische wohlgeformt sind):

1. Dem Ausdruck $\neg X$ wird der Wahrheitswert F bzw. W zugeordnet, je nachdem ob der Ausdruck X den Wahrheitswert W bzw. F besitzt.
2. Dem Ausdruck $(X \wedge Y)$ wird der Wahrheitswert W zugeordnet, falls sowohl X als auch Y den Wahrheitswert W aufweisen, andernfalls erhält der Ausdruck den Wert F .
3. Dem Ausdruck $(X \vee Y)$ wird der Wahrheitswert W zugeordnet, falls entweder X oder Y oder beide den Wahrheitswert W aufweisen, andernfalls erhält der Ausdruck den Wert F .
4. Dem Ausdruck $(X \rightarrow Y)$ wird der Wahrheitswert W zugeordnet, falls entweder X den Wert F oder Y den Wert W aufweist, andernfalls erhält der Ausdruck den Wert F .
5. Dem Ausdruck $(X \leftrightarrow Y)$ wird der Wahrheitswert W zugeordnet, falls X und Y den gleichen Wahrheitswert besitzen, andernfalls erhält der Ausdruck den Wert F .

Das Beispiel demonstriert, wie eine Sprache aufgebaut wird:

1. Es wird der *syntaktische Teil* spezifiziert: Dieser besteht aus der Spezifikation der Grundausrücke und der syntaktischen Regeln.

2. Anschliessend folgt die Spezifikation des *semantischen Teils*, nämlich, die semantische Interpretation der Grundausdrücke und der semantischen Regeln, welche die Bedeutung komplexerer Ausdrücke aufgrund der Bedeutungen der elementarerer festlegen.

Der Begriff des *Modells* bezieht sich immer auf den semantischen Teil. Im Falle der Aussagenlogik bilden die Wahrheitswertbelegungen der Grundausdrücke die Modelle. Man kann ein Modell daher einfach als eine Sequenz von Wahrheitswerten betrachten. Diese Sequenz repräsentiert eine konkrete Instanz, d.h. eine konkrete Implementation des Ausdrucks innerhalb der »Welt der Wahrheitswerte«.

Die oben gegebene Formel $((\neg A \vee B) \leftrightarrow (A \rightarrow B))$ ist *unter allen Modellen* (Wahrheitswertbelegungen) *wahr*. Damit ist ihre Negation *unter allen Modellen falsch*.

Ist eine Formel unter allen Modellen wahr, so nennt man sie *notwendig wahr* oder auch eine *Tautologie*.

Ist eine Formel unter allen Modellen falsch, so nennt man sie *notwendig falsch* oder auch eine *Kontradiktion*.

Ist eine Formel unter einigen Modellen wahr und unter anderen falsch, so ist sie *möglicherweise wahr* oder *möglicherweise falsch*.

Kehren wir nach dieser kurzen Exkursion über Syntax und Semantik in der Aussagenlogik und die Beutung von Modellen innerhalb dieses Systems wieder zu unseren partiellen potentiellen Modellen zurück. Die in M_{pp} gegebene Beschreibung repräsentiert alle möglichen Instanzen innerhalb der Welt, auf welche die Theorie angewendet wird. In Bsp.2-47 sind dies alle Entitäten, die mit der Stimulumsmenge S konfrontiert sind, und die Responsemenge R zur Verfügung haben. Durch wiederholte Präsentation der Stimulumsmenge und der Evokation von Responses ergeben sich relative Häufigkeiten, die als Schätzwerte für die in P spezifizierten Wahrscheinlichkeiten betrachtet werden können.

Im aktuellen Beispiel kann es sich bei einem Modell um einen Menschen, ein Tier, aber auch um eine Maschine oder ein Testverfahren handeln. Die Antworten »A« und »N« können hierbei durch Drücken einer Taste oder durch Aufleuchten eines Lichts (im Falle einer Maschine) gegeben werden.

Der zentrale Aspekt der M_{pp} besteht in der Tatsache, dass sie das *Universum aller möglichen Anwendungen des Theoriekerns* repräsentiert. Ein Modell, das nicht durch die in M_{pp} gegebene Spezifikation charakterisiert werden kann, liegt ausserhalb der möglichen Anwendungen des aktuellen Theoriekerns. Auf diese Weise wird durch die Theorie selbst die Menge aller möglichen Anwendung vorgegeben, *Die Welt*

wird sozusagen durch die Brille der Theorie betrachtet, indem sie die Menge der möglichen Anwendungen selbst bestimmt.

Wie oben ausgeführt, kann die Menge M_{pp} allgemeiner charakterisiert werden, sodass sie viel mehr mögliche Anwendungen enthält.



Bsp.2-49: Verallgemeinerung der Spezifikation der Menge partiellen Modelle:

Die Verallgemeinerung der gegebenen Spezifikation von M_{pp} kann in zwei möglichen Erweiterungen bestehen:

1. Die Anzahl der Stimuli kann beliebig sein, ebenso die Anzahl der Responses. Die Anzahl der Stimuli muss nicht mit der Anzahl der Responses übereinstimmen.
2. Die Stimuli können mehrdimensional sein. Dies ist für die meisten realistischen Stimuli der Fall. Ein Stimulus S_i ($i = 1, 2, \dots, n$) wird daher als ein Vektor $S_i = [S_{i1}, S_{i2}, \dots, S_{is_i}]$ betrachtet, wobei die S_{ik} ($k = 1, 2, \dots, s_i$) die Werte der Stimuluskomponenten auf verschiedenen Dimensionen bezeichnen.

Analog können die Responses mehrdimensional sein:

$$R_j = [R_{j1}, R_{j2}, \dots, R_{jr_j}] \quad (j = 1, 2, \dots, m).$$



Bemerkung zur Notation:

Als Symbole für Vektoren und Matrizen, werden fett gedruckte Buchstaben verwendet.

Die Anzahl der Dimensionen der Stimulus- und Responsevektoren kann zwischen den Stimuli (bzw. Responses) variieren, daher sind die Bezeichnungen s_i bzw. r_j für die Anzahl Komponenten vom Index i bzw. j abhängig.

Durch diese Erweiterung erhöht sich die Menge der möglichen Anwendungen der Theorie. Betrachten wir als nächstes die Menge M_p der potentiellen Modelle, die sich durch Einführung der T-theoretischen Begriffe aus der Menge M_{pp} ergibt.

2.3.3.2 DIE MENGE DER POTENTIELLEN MODELLE EINER THEORIE

Die Menge der potentiellen Modelle M_p ist in der Menge partiell potentiellen Modelle M_{pp} eingebettet, d.h. M_p ist eine Teilmenge von M_{pp} . Dies ergibt sich, indem jedes Element von M_p als zusätzliche Erfordernis die T-theoretischen Komponenten erfüllen muss.

Im aktuellen partiellen Modell M_p von Bsp.2-47 bedeutet dies, dass ein Element (Mensch, Tier, Maschine oder Verfahren) nicht nur der durch M_{pp} beschriebenen Situation ausgesetzt ist, sondern – zusätzlich – über eine latente Entscheidungsdimension θ , mit zugehörigen Gauss'schen Verteilungsfunktionen Φ_N und Φ_A , sowie einer latenten Entscheidungsfunktion δ verfügt.

Bei diesen Grössen handelt es sich um T-theoretische Grössen. Hier stellt sich nun unmittelbar die folgende Frage:



Frage:

Was genau bedeutet T-theoretisch und wie unterscheiden sich T-theoretische Begriffe von theoretischen?

Zur Klärung dieser Frage betrachten wir zuerst den Begriff der *T-theoretischen Messung*.



Konzept 2-14: T-theoretische Messung:

Die Messung einer Grösse ist *T-theoretisch*, falls zur Ermittlung der Werte der Grösse, die Gültigkeit der Theorie *T* vorausgesetzt wird.

Man beachte, dass diese Charakterisierung exakt dem Konzept der theorieabhängigen Messung entspricht (vgl. Konzept 2-7). Beide Konzepte sind also identisch. Aufgrund dieser Definition ergibt sich nun das Konzept der *T-theoretischen Grösse* wie folgt.



Konzept 2-15: T-theoretische Grössen:

Eine Grösse ist *T-theoretisch*, falls sie in allen Anwendungen der Theorie *T* nur T-theoretisch gemessen werden kann.

Im konkreten Fall können die Parameter der Verteilungsfunktionen Φ_N , Φ_A und der Entscheidungsfunktion δ in allen Anwendungen der Signalentdeckungstheorie (SDT) nur theorieabhängig (oder *SDT-theoretisch*) gemessen werden. Daher handelt es sich um *SDT-theoretische* Grössen. Die konkreten Werte der latenten θ Entscheidungsdimension werden nicht direkt gemessen (nur deren Verteilung). Da jedoch die anderen SDT-theoretischen Grössen auf der Existenz dieses Konstrukts beruhen, handelt es sich bei θ ebenfalls um ein SDT-theoretisches Konstrukt, das nur innerhalb der SDT eine Bedeutung hat.



Bemerkung zur Bedeutung latenter Variablen in psychologischen Modellen:

Viele psychologische Modelle umfassen latente Variablen – ähnliche wie die latente Entscheidungsvariable θ innerhalb der SDT-Modelle. Die wichtigsten Modelle mit dieser Eigenschaft sind:

- Faktorenanalytische Modelle (Bollen, 1989);
- Multinomiale Prozessbaummodelle (Batchelder & Riefer, 1999) [siehe Bsp.2-21];
- Item-Response Modelle (Lord & Novick, 1968);
- Latente Klassenmodelle (Clogg, 1995);
- Thurstone-Modelle (Thurstone, 1927).

Nach dieser Klärung des Begriffs der T-theoretischen Grösse, kommen wir zum zweiten Teil der oben gestellten Frage: Wie unterscheiden sich T-theoretische von theoretischen Grössen?

Der zentrale Unterschied zwischen theoretischen und T-theoretischen Grössen besteht darin, dass sich letztere *immer auf eine bestimmte Theorie beziehen*. Diese Unterscheidung ist von zentraler Bedeutung, denn aufgrund dieser Eigenschaft entgehen T-theoretische Terme jener Kritik, welchen theoretische Terme ausgesetzt sind. Wie bereits oben ausgeführt wurde, ist die Unterscheidung zwischen einer Beobachtungssprache und einer theoretischen Sprache, wie sie von den Empiristen konzipiert wurde (vgl. Prinzip 2-2), künstlicher Natur, da es sich eher um ein Kontinuum handelt (vgl. Prinzip 2-3).

Dieses Problem besteht für T-theoretische Terme nicht. Hier gibt es ein eindeutiges Kriterium, das festlegt, ob eine Grösse T-theoretisch ist, nämlich, ob die Theorie als wahr vorausgesetzt werden muss, um die T-theoretische Grösse zu messen (in allen Anwendungen der Theorie).

Man beachte auch den folgenden wichtigen Zusammenhang: T-nicht-theoretische Grössen können T-theoretisch sein bezüglich einer anderen Theorie.



Bsp.2-50: *SDT-nicht-theoretische Grössen als T-theoretische Grössen bezüglich einer anderen Theorie:*

Gegeben:

Die Stimulussmenge S bestehe aus Gesichtern, die bezüglich ihres emotionalen Ausdrucks zu unterscheiden sind.

Der emotionale Ausdruck der gezeigten Gesichter wurde mit Hilfe des facial action coding systems von Ekman und Friesen (1978) erzeugt.

In diesem Falle ist S SDT-nicht-theoretisch, jedoch T-theoretisch bezüglich der Theorie von Ekman und Friesen, auf der das Kodierungssystem beruht.

Wir haben damit die zentrale Bedeutung des Konzepts der T-theoretischen Grössen und die Beziehung zwischen potentiell partiellen und partiellen Modellen erklärt. Zu erwähnen ist noch, dass – ähnlich wie im Falle der T-nicht-theoretischen Terme – abstraktere, generellere oder auch einfach alternative Spezifikationen der T-theoretischen Grössen möglich sind. So kann für bestimmte Anwendungen die latente Entscheidungsvariable als mehrdimensional betrachtet werden. Ebenso können alternative Entscheidungs- und Verteilungsfunktionen angenommen werden.

Die Spezifikation der Menge M_p der potentiellen Modelle enthält alle Konstrukte der Theorie bzw. der konkreten Anwendung der Theorie. In der Spezifikation von M_p sind jedoch noch nicht die gesetzesmässigen Annahmen der Theorie spezifiziert. Dies geschieht, indem M_p zur Menge der Modelle M vervollständigt wird.

2.3.3.3 DIE MENGE DER MODELLE EINER THEORIE

Im Falle der SDT bestehen die Gesetzesaussagen der Theorie in quantitativen Aussagen über die Wahrscheinlichkeiten der unterschiedlichen bedingten Wahrscheinlichkeiten der möglichen Antworten bei gegebenem Stimulus. Diese Aussagen sind in der Spezifikation von M enthalten (Gleichung (2-9)), zusätzlich zu den in M_p enthaltenen Spezifikationen. Daher ist die Menge M eine (echte) Teilmenge von M_p : Jedes Modell von M ist auch ein Modell von M_p , aber nicht umgekehrt.

Das oben spezifizierte Gauss'sche SDT-Modell sagt aus, dass die bedingten Antwortwahrscheinlichkeiten durch die in (2-9) spezifizierten Gleichungen gegeben sind. Die vom Modell vorhergesagten Wahrscheinlichkeiten entsprechen daher unterschiedlichen Flächen unter den Normalverteilungskurven.

Diese Gesetzesaussagen bilden den empirisch prüfbaren Teil, indem die Vorhersagen des Modells mit den beobachteten Wahrscheinlichkeiten, welche in M_{pp} spezifiziert wurden, in Beziehung gesetzt werden. Falls eine Person, ein Tier, eine Maschine oder ein Testverfahren nicht nur ein partielles Modell M_p ist, sondern zusätzlich ein Modell M , dann sollten die beobachteten bedingten Wahrscheinlichkeiten P den vom Modell vorhergesagten theoretischen Wahrscheinlichkeiten π – bis auf Zufallsabweichungen, die sich aufgrund der beschränkten Stichprobe ergeben – entsprechen.



Bemerkung zur Prüfbarkeit von Modellen mit freien Parametern:

Das SDT-Modell besitzt freie Parameter, die aus den Daten geschätzt werden. Hierbei handelt es sich um:

- Die Entscheidungsschwelle τ .
- Den Mittelwerts- und Varianzparameter der Verteilungsfunktion für neue Items: μ_N, σ_N^2 .
- Den Mittelwerts- und Varianzparameter der Verteilungsfunktion für alte Items: μ_A, σ_A^2 .

Diese Parameter werden aus den Daten geschätzt. Das Modell ist im Allgemeinen nur dann testbar, wenn die Anzahl der freien Parameter kleiner ist als die Anzahl Datenpunkte.

Im aktuellen Beispiel gibt es 2 freie Datenpunkte, beispielsweise $P(\gg N \ll N)$ und $P(\gg N \ll A)$. Aufgrund der Beziehungen $P(\gg A \ll N) = 1 - P(\gg N \ll N)$ und $P(\gg A \ll A) = 1 - P(\gg N \ll A)$ sind die anderen beiden Wahrscheinlichkeiten völlig von den beiden gewählten abhängig.

Das Modell ist daher nur dann prüfbar, wenn mindestens 4 der 5 freien Parameter auf bestimmte Werte fixiert werden.

Die Formulierung des Modells illustriert auch das Problem des Holismus, denn aufgrund der gegebenen Spezifikation kann bei falschen Vorhersagen des Modells bei einer gegebenen Anwendung nicht bestimmt werden, ob die Verteilungen nicht die korrekt Form haben, oder ob die Entscheidungsfunktion falsch spezifiziert wurde. Beide Komponenten sind nämlich für die Bestimmung der modellbasierten Vorhersagen notwendig. Allerdings ist es möglich, die einzelnen Komponenten – zumindest indirekt – zu testen, indem experimentelle Manipulationen durchgeführt werden, welche (hauptsächlich) nur eine der beiden Komponenten betreffen. Es lässt sich dann mit Hilfe eines erweiterten Modells, welche zusätzlich Gesetzaussagen über den Effekt der Manipulationen auf Entscheidungs- und Verteilungsfunktionen enthält, prüfen, ob die Manipulationen den Vorhersagen des erweiterten Modells entsprechen.

Damit haben wir die 3 wichtigsten Komponenten des Theoriekerns besprochen. Es gibt jedoch noch zwei weitere Komponenten, welche im Folgenden kurz behandelt werden.

2.3.3.4 BESCHRÄNKUNGEN IN PARTIELLEN MODELLEN UND VERBINDUNGEN ZWISCHEN PARTIELLEN MODELLEN

Neben den drei Arten von Modellen M_{pp} , M_p und M umfasst der Theoriekern noch zwei weitere Komponenten: Beschränkungen in partiellen Modellen und Verbindungen zwischen partiellen Modellen aus unterschiedlichen Anwendungen.

Ein wichtiges Beispiel für Beschränkungen bezüglich der aktuell diskutierten Menge der partiellen Modelle hatten wir bereits besprochen, nämlich die Fixierung von Parametern der Verteilungsfunktionen.



Bsp.2-51: Fixieren von Parametern im SDT-Modell:

Gegeben:

Das SDT-Modell von Bsp.2-47.

Damit das Modell überhaupt aus den Daten geschätzt werden kann, müssen mindestens 3 Parameter fixiert werden.

Üblicherweise werden die folgenden Parameter wie folgt fixiert:

- $\mu_N = 0$
- $\sigma_N^2 = \sigma_A^2 = 1$

Damit kann dann der Schwellenwert τ und der Mittelwert μ_A aus den Daten geschätzt werden. Ersterer gibt darüber Auskunft, ob eine Responsebias vorliegt, d.h. ob die Person, Maschine etc. bestimmte Antworten bevorzugt. Dies ist dann der Fall, wenn gilt: $\tau \neq \mu_A/2$, d.h. die Schwelle liegt nicht genau in der Mitte zwischen den beiden Verteilungen. μ_A gibt Auskunft über die Sensitivität des Entscheiders: Je grösser μ_A , desto weiter liegen die beiden Normalverteilungskurven auseinander und desto sensibler ist daher der Entscheider.

Das Modell ist damit noch nicht testbar. Man kann jedoch verschiedene Hypothesen testen. Beispielsweise lässt sich die Hypothese prüfen, dass der Entscheider keine Responsebias hat. Hierzu setzt man $\tau = \mu_A/2$ und vergleicht die Modellvorhersagen mit den Daten. Falls die Vorhersagen nicht signifikant von den beobachteten Werten abweichen, so kann man annehmen, dass der Entscheider keine Bevorzugung für bestimmte Antworten aufweist.

Man beachte jedoch, dass dieser Test die Korrektheit der Modellannahmen voraussetzt.

Beschränkungen können auch zwischen partiellen Modellen aus verschiedenen Anwendungen bestehen.



Bsp.2-52: *Beschränkungen zwischen zwei Anwendungen:*

Gegeben:

Zwei Anwendungen des SDT-Modell von Bsp.2-47 mit zwei Stimulismengen:

- $S_{\text{schwach}} = \{N, A_{\text{schwach}}\}$ sei die Menge der Stimuli, mit schwachen alten Items, d.h. die alten Items sind schwer von den neuen zu unterscheiden.
- $S_{\text{stark}} = \{N, A_{\text{stark}}\}$ sei die Menge der Stimuli, mit starken alten Items, d.h. die alten Items sind gut von den neuen zu unterscheiden.

Eine mögliche Beschränkung zwischen den beiden Anwendungen bestünde in der Annahme, dass die Schwellenparameter τ für beide Anwendungen identisch sind.

Die Einführung von Beschränkungen macht in vielen Fällen ein Modell bzw. eine Theorie erst testbar bzw. sie erhöht den Grad der Prüfbarkeit.

Ein weiterer Aspekt betrifft die Verbindungen zwischen unterschiedlichen partiellen Modellen aus unterschiedlichen Theorien. So könnte man beispielsweise globale Verbindungen zu den globalen Memory Matching Modellen (Clark & Gronlund, 1996) für das Wiedererkennen herstellen, wobei die Vertrautheit im SDT-Modell mit dem Grad der Aktivierung von Gedächtniseinheiten in Beziehung gesetzt wird.

Mit diesen Ausführungen zu globalen Beschränkungen und Verbindungen beschliessen wir die Darstellung der Struktur des Theoriekerns.

2.3.3.5 INTENDIERTE ANWENDUNGEN

Die zweite Komponente eines Theorieelements (Konzept 2-11) betrifft die *Menge der intendierten Anwendungen I*. Hierbei handelt es sich um eine Teilmenge der partiellen potentiellen Modelle M_{pp} , welche die verschiedenen empirischen Anwendungen der Theorie umfasst. Es lassen sich vier Arten von intendierten Anwendungen unterscheiden:

1. *Paradigmatische Anwendungen*: Hier handelt es sich um jene Anwendungen, für welche die Theorie geschaffen wurde.
2. *Erfolgreiche Anwendungen*: Dies sind erfolgreiche Neuanwendungen, die sich im Zuge des Ausbaus des Forschungsprogramms ergeben haben.
3. *Vermutete Anwendungen*: Hier handelt es sich um mögliche Anwendungen, von denen angenommen wird, dass sie erfolgreich sein werden, deren Erfolg aber noch nicht feststeht.
4. *Zweifelhafte Anwendungen*: Hierbei handelt es sich um Anwendungen, die sich als (bisher) nicht erfolgreich herausgestellt haben.

Die Vertreter des Ansatzes gehen davon aus, dass bei Scheitern einer Anwendung, diese aus der Theorie der intendierten Anwendung entfernt werden kann, ohne dass dies der Theorie wesentlich schadet.



Bsp.2-53: *Entfernung von Anwendung aus der Menge der intendierten Anwendungen:*

Ein klassisches Beispiel für die Elimination einer Anwendung aus der Menge der intendierten Anwendung betrifft die Anwendung der Theorie Newtons auf Phänomene des Lichts.

Nachdem klar geworden war, dass die Theorie Newtons verschiedene das Licht betreffende Phänomene nicht erklären kann, wurden diese Anwendungen aus der Menge der intendierten Anwendungen entfernt (Stegmüller, 1985).

Ein vergleichbares Beispiel findet sich in der Psychologie: Nachdem klar wurde, dass die meisten Formen menschlichen Lernens durch assoziative Lerntheorien nicht erklärt werden können, wurden diese aus der Menge der intendierten Anwendungen entfernt.

Damit beschliessen wir die Darstellung der strukturalistischen Konzeption der Wissenschaft. Bevor wir das Thema verlassen, seien noch die Vor- und Nachteile des Ansatzes kurz diskutiert.

2.3.3.6 WÜRDIGUNG: STÄRKEN UND SCHWÄCHEN DES ANSATZES

Die strukturalistische Konzeption weist die folgenden positiven Eigenschaften auf:

- Alle Komponenten einer Theorie werden explizit gemacht: Dies betrifft nicht nur die Gesetze sondern auch die Entitäten, welche die

Theorie voraussetzt. Damit erfüllt der Ansatz in idealer Weise das Prinzip, »alle Annahmen auf den Tisch zu legen« und so der Kritik zugänglich zu machen (Prinzip 2-5).

- Die Unterscheidung zwischen T-theoretischen und T-nicht-theoretischen Konstrukten ist wichtig, weil sie klar macht, welche Entitäten spezifisch für die vorliegende Theorie sind.
- Der Ansatz ermöglicht es in idealer Weise Verbindungen zwischen Theorieelementen (Erweiterungen, Modifikationen, Spezialisierungen) sowie Beschränkungen zwischen diesen explizit zu machen. Auf diese Weise kann das gesamte Netz von Theorien rekonstruiert bzw. verdeutlicht werden.
- Der Ansatz erlaubt es auch Entwicklungen von Theorien zu rekonstruieren.

Diesen Vorteilen des Ansatzes stehen mindestens drei Nachteile gegenüber:

- Der Ansatz bringt eine »Aufblähung der Notation« mit sich. Damit ist gemeint, dass schon für die Rekonstruktion ganz einfacher Theorien ein gewaltiger Satz an Symbolen und Notationen benötigt werden. Dies macht den Ansatz extrem schwerfällig, obwohl durch die informelle mengentheoretische Darstellung, die Rekonstruktionen im Prinzip nicht schwer zu verstehen sind. Dennoch verhindert diese »Aufblähung der Notation« eine grosse Verbreitung des Ansatzes (Dies trotz der heroischen Bemühungen der Autoren in Westmeyer, 1989 und 1992).
- Das Problem des Entfernens von Anwendungen aus der Menge *I* der intendierten Anwendungen scheint mir irreführend dargestellt. So schreibt z.B. Stegmüller (1985, Seite, 23):

Wenn die Theorie für ein Element von I vollkommen versagt, so wird dieses Element einfach aus der Menge der intendierten Anwendungen I entfernt. Wer darin eine ‚Immunsierungsstrategie‘ erblickt, der huldigt nicht einem kritischen sondern einem überspannten Rationalismus (Stegmüller, 1985, Seite, 23).

Diese Auffassung trifft m. E. die Sache nicht ganz, denn manchmal kann man Anwendungen nicht einfach aus der Menge der intendierten Anwendungen entfernen ohne die Theorie wesentlich zu schwächen. So kann z.B. die Berechnung der Eigendrehung des Merkur nicht aufgrund der beobachteten Anomalien aus der Menge der intendierten Anwendungen der Newtonschen Partikelmechanik entfernt werden, da dies zu deren ureigenstem Bereich gehört. Ebenso wenig kann kausales Lernen aus der Menge der intendierten Anwendungen der assoziativen Lerntheorie entfernt werden, ohne diese empfindlich zu schwächen.

- Der Ansatz liefert keine Kriterien um wissenschaftliche von pseudo- oder protowissenschaftlichen Theorien zu unterscheiden. Dies zeigt sich m. E. am besten darin, dass es zahlreiche Versuche gibt

die Freudsche Neurosentheorie mit Hilfe des Ansatzes zu rekonstruieren (z.B. Balzer & Marcou, 1989). Ebenso gut könnte man versuchen, Theorien der Alchemie und Astrologie mit Hilfe des Ansatzes »rational« zu rekonstruieren.

Trotz dieser Kritikpunkte muss die grosse Bedeutung des Ansatzes für die Rekonstruktion und den Vergleich von Theorien betont werden.

2.4 Übungen zu Kapitel 2



Übung 2-1:

1. Geben Sie Beispiele für folgende deduktive Schlussfolgerungen:
 - Modus Ponens
 - Modus Tollens
 - Fehlerhaft: Bejahung der Konsequenz
 - Fehlerhaft: Verneinung des Antezedens.
 2. Präsentieren Sie ein Beispiel zur Explikation der Konzepte der hinreichenden und notwendigen Bedingung
- Geben sie Beispiele für folgende induktive Schlussfolgerungen:
- Statistische Verallgemeinerung
 - Analogieschluss
3. Geben sie ein Beispiel für eine abduktive Schlussfolgerung:



Übung 2-2:

Geben Sie zwei eigene Beispiele von modellabhängiger Messung und zeigen Sie, wie die Messungen zur Prüfung von Theorien eingesetzt werden können.

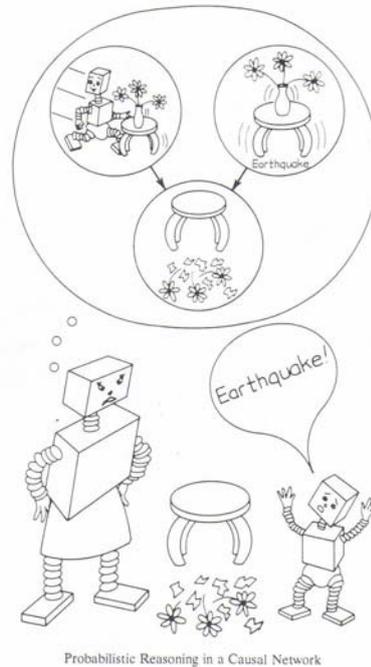
**Übung 2-3:**

Abb. 2.12 Kausales Schlussfolgern in Bayes-Netzen (Neapolitan, 1990).

Erklären Sie, welches Prinzip durch Abb. 2.12 expliziert werden soll.

**Übung 2-4:**

Welchem Prinzip von Popper widerspricht die Annahme der Theologie als Wissenschaft?

**Übung 2-5:**

Zeigen Sie mit Hilfe einer Wahrheitstafel die Äquivalenz der beiden logischen Ausdrücke: $A \wedge B$ und $\neg(\neg A \vee \neg B)$.

**Übung 2-6:**

Diskutieren Sie ein Beispiel für Unterbestimmtheit, bzw. für das Vorliegen von Alternativerklärungen empirischer Befunde in Ihrem Fachgebiet.

**Übung 2-7:**

Geben Sie ein Beispiel für ein Forschungsprogramm aus dem Gebiet der Psychologie (oder des von Ihnen studierten Faches). Wie sieht der Theoriekern aus, welche Zusatzannahmen werden gemacht?

**Übung 2-8:**

Spezifizieren Sie die relevanten Komponenten der partiellen Modells M_p für das Prozessbaummodell zur Messung von Wissen und Raten aus Bsp.2-21.

3. Wissenschaftliche Erklärungen in der Psychologie

In diesem Kapitel wird die Frage behandelt, was eine gute wissenschaftliche psychologische Erklärung ausmacht. Im ersten Abschnitt wird die allgemeine Struktur wissenschaftlicher Erklärungen erläutert. Es folgt eine Darstellung der Idee von David Marr (1945–1980) zu den verschiedenen Erklärungsebenen eines Informationsverarbeitungssystems. Im dritten Abschnitt wird das Problem der Reduktion psychologischer Erklärungen behandelt. Im letzten Abschnitt des Kapitels wird kurz das Problem von Pseudoerklärungen in der Psychologie besprochen.

3.1 Die Struktur wissenschaftlicher Erklärungen

Der wohl einflussreichste Versuch zur Aufdeckung der Struktur wissenschaftlicher Erklärungen stammt von Carl Gustav Hempel (1905–1997) und Paul Oppenheim (1886–1977) aus dem Jahr 1948. Obwohl sich heute die meisten Wissenschaftstheoretiker darüber einig sind, dass dieser Versuch missglückt ist (Curd & Cover, 1998; Ruben, 1993; Salmon, 1989, 2006), findet man diesen Ansatz in praktisch allen Abhandlungen zur wissenschaftlichen Erklärung. Wir beginnen daher mit einem kurzen Abriss dieses Ansatzes.

3.1.1 Die »Grossmutter« der Struktur wissenschaftlicher Erklärungen: Das Hempel-Oppenheim-Schema

Gemäss der Konzeption von Hempel und Oppenheim besteht eine wissenschaftliche Erklärung in einer logischen Schlussfolgerung. Das Hempel-Oppenheim-Schema (HO-Schema) präsentiert sich in zwei grundsätzlichen Varianten, der *deduktiv-nomologischen* und der *induktiv-probabilistischen* Form. Wir behandeln zuerst diese beiden Grund-schemata und untersuchen dann die Probleme des gesamten Ansatzes.



Bemerkung:

Es wird hier nur das grundlegende Schema zur Erklärung von Einzelereignissen behandelt. Die Anwendung des Schemas zur Ableitung von Gesetzhypothese aus Theorien und Nebenbedingungen wird nicht besprochen.

Weiters gibt es verschiedene technische Details, welche notwendig sind, um bestimmte triviale Schlussfolgerungen auszuschliessen. Auch hierauf wird im Folgenden nicht eingegangen. Details hierzu findet man in der oben angegebenen Literatur.

3.1.1.1 DAS HEMPEL-OPPENHEIM-SCHEMA DER ERKLÄRUNG: DE- DUKTIV-NOMOLOGISCHE FORM



Konzept 3-1: *Das Hempel-Oppenheim-Schema (HO-Schema) der Erklärung: Die deduktiv-nomologische Version:*

Gegeben:

1. Das *Explanandum* E , d.h. das zu erklärende Phänomen,
2. Eine Menge von *Gesetzesaussagen*: G_1, G_2, \dots, G_n , die für die Erklärung notwendig sind (d.h. ohne Gesetzesaussagen kann keine Folgerung durchgeführt werden).
3. *Antezedensbedingungen*: A_1, A_2, \dots, A_m , welche die Voraussetzungen für die Anwendung der Gesetzesaussagen festlegen.

Eine wissenschaftliche Erklärung besteht in der *logischen Ableitung* des Explanandums aus den Gesetzesaussagen und den Antezedensbedingungen:

$$\frac{G_1, G_2, \dots, G_n}{A_1, A_2, \dots, A_m} \\ E$$



Bemerkung zur Notation:

Die horizontale Linie symbolisiert die Schlussfolgerung. Sie hat die gleiche Bedeutung wie das oben verwendete Symbol \therefore .

Ein einfaches Beispiel illustriert die Funktionsweise des HO-Schemas:



Bsp.3-1: *Erklärung eines emotionalen Zustandes (HO-Schema deduktive Version):*

Explanandum (E):

Hans hat sich gestern über Karin geärgert.

Antezedensbedingung (A):

Karin hat Hans ohne Angabe von Gründen versetzt.

Gesetzesaussage (G):

Wenn eine Person X von einer anderen Person Y versetzt wird, so wird X sich über Y ärgern, falls Y keine Gründe für ihr Verhalten angeben kann.

Um zu zeigen, dass es sich um eine korrekte Schlussfolgerung handelt, kann man die Aussagen vollständig formalisieren. Die Schlussfolgerung sieht dann wie folgt aus:

G: $\forall x \forall y [\text{VERSETZT}(x, y) \wedge \neg \text{GRUND}(x) \rightarrow \text{ÄRGER}(y, x)]$

A: $\text{VERSETZT}(k, h) \wedge \neg \text{GRUND}(k)$

E: ÄRGER(h, k)

Die Symbole haben die folgende Bedeutung:

VERSETZT(k, h)	Karin hat Hans versetzt
¬GRUND(k)	Karin hat keine Begründung für ihr Verhalten angegeben.
ÄRGER(h, k)	Hans ärgert sich über Karin
∧	symbolisiert das logische UND
¬	symbolisiert das logische NICHT
→	symbolisiert die logische Implikation
∀x	symbolisiert die Tatsache, dass für alle Objekte x aus einem vorgegebenen Bereich die nachfolgende Formel zutrifft.

Die logische Struktur einer Erklärung entspricht einem *Modus Ponens mit Instantiierung*. Das bekannteste Beispiel für eine derartige Schlussfolgerung lautet:

Alle Menschen sind sterblich.	$\forall x$ [MENSCH(x) → STERBLICH(x)]
Sokrates ist ein Mensch.	MENSCH(s)
Sokrates ist sterblich.	STERBLICH(s)

3.1.1.2 DAS HEMPEL-OPPENHEIM–SCHEMA DER ERKLÄRUNG: INDUKTIV-PROBABILISTISCHE VERSION

Die deduktive Form des HO-Schemas hat einen gewaltigen Nachteil: Speziell im Bereich der Sozialwissenschaften hat man ganz selten deterministische Gesetzmässigkeiten. Meistens besteht nur ein probabilistischer Zusammenhang zwischen Ereignissen und die Gesetze sind daher probabilistischer Natur.

Die oben formulierte Gesetzesaussage würde demgemäss wie folgt aussehen:



Bsp.3-2: Probabilistische Gesetzesaussage:

Die oben dargestellte Gesetzesaussage in probabilistischer Form sieht wie folgt aus:

Wenn eine Person X von einer anderen Person Y versetzt wird, so wird sich X über Y mit Wahrscheinlichkeit p ärgern, falls Y keine Gründe für ihr Verhalten angeben kann (wobei p ein Wert möglichst nahe bei 1.0 sein sollte).

Bei Verwendung probabilistischer Gesetze sieht das HO-Schema wie folgt aus:



Konzept 3-2: Das Hempel-Oppenheim-Schema (HO-Schema) der Erklärung: Induktiv-probabilistische Version:

Gegeben:

1. Das *Explanandum* E , d.h. das zu erklärende Phänomen,
2. Eine Menge von *Gesetzesaussagen*: G_1, G_2, \dots, G_n ,
3. *Antezedensbedingungen*: A_1, A_2, \dots, A_m , welche die Voraussetzungen für die Anwendung der Gesetzesaussagen festlegen.

Eine wissenschaftliche Erklärung besteht nun in der induktiven Schlussfolgerung:

$$\frac{G_1, G_2, \dots, G_n}{A_1, A_2, \dots, A_m} \quad p \quad \text{—} \\ E$$



Bemerkung zur Notation:

Die horizontale Linie mit dem Symbol p symbolisiert, dass das Explanandum E mit Wahrscheinlichkeit p aus den Prämissen folgt, wobei p möglichst gross sein soll.



Bsp.3-3: Erklärung eines emotionalen Zustandes (induktive Form):

G: $\forall x \forall y [\text{VERSETZT}(x, y) \wedge \neg \text{GRUND}(x) \xrightarrow{p} \text{ÄRGER}(y, x)]$

A: $\text{VERSETZT}(k, h) \wedge \neg \text{GRUND}(k)$ p —————

E: $\text{ÄRGER}(h, k)$

Die Symbole haben folgende Bedeutung

\xrightarrow{p} Symbolisiert eine probabilistische Implikation:

Falls der Vordersatz wahr ist, so ist mit Wahrscheinlichkeit p auch der Nachsatz wahr. Dies bedeutet, dass bei Gültigkeit der Implikation die Korrektheit des Vordersatzes nicht mehr notwendig die Wahrheit des Nachsatzes nach sich zieht.

Die restlichen Symbole haben die gleiche Bedeutung wie oben.

In Worten ausgedrückt lautet die obige Erklärung wie folgt:

Wenn eine Person X von einer anderen Person Y versetzt wird, so wird sich mit hoher Wahrscheinlichkeit X über Y ärgern, falls Y keine Gründe für ihr Verhalten angeben kann.

A: Karin hat Hans ohne Angabe von Gründen versetzt.

_____ *sehr wahrscheinlich* –
E: Hans ärgert sich über Karin.

Für die induktiv-probabilistische Version des HO-Schemas ergibt sich eine Schwierigkeit, die für die deduktiv-nomologische Version nicht existiert:

Es ist möglich, dass unsere Person Hans eine andere Eigenschaft besitzt, die es sehr unwahrscheinlich macht, dass er sich über das Verhalten von Karin ärgert, z.B. könnte Hans – immer wenn er versetzt wird – 50 mg Valium einnehmen, was dazu führt, dass ihm alles egal ist und er sich nicht ärgert.

Hempel hat dieses Problem erkannt und als Lösung vorgeschlagen, dass bei induktiven Schlussfolgerungen *das gesamte verfügbare Erfahrungswissen* einbezogen werden sollte, insofern es für die aktuelle Schlussfolgerung relevant ist. Damit liesse sich eine Differenzierung für alle möglichen Kombinationen von Merkmalen und den relevanten Folgerungen finden.

Diese Lösung hat zwei grosse Probleme:

- (1) Sie benötigt die Annahme einer geschlossenen Welt, d.h. damit diese Lösung tatsächlich funktioniert, ist das Erfahrungswissen als abgeschlossen zu betrachten. Andernfalls können – aufgrund von neuen Differenzierungen – die bisher gemachten Schlüsse falsch sein.
- (2) Es gibt eine Explosion der Komplexität, da für jede Beurteilung, ob eine bestimmte Prämisse anwendbar sei, extrem viele Spezialfälle beachtet werden müssen.



Bemerkung:

Der Grund, warum für die deduktive-nomologische Form dieses Problem nicht existiert, lautet wie folgt:

Falls $A \rightarrow B$ gilt, so gilt auch: $(A \wedge C_1 \wedge C_2 \wedge \dots \wedge C_n) \rightarrow B$, unabhängig vom Wahrheitswert der Aussagen C_1, C_2, \dots, C_n .

Folgt daher eine Aussage B aus einer anderen A , so ändert das Hinzufügen weiterer Prämissen ändert nichts mehr daran, dass bei Gültigkeit von A auf B geschlossen werden kann.

Die Tatsache, dass das Auftauchen eines neuen Faktums die Wahrscheinlichkeit eines Ereignisses bzw. eines Zusammenhanges völlig verändern kann, wird eindrucksvoll durch das so genannte *Simpson-Paradoxon* (Simpson, 1951) belegt, das aufgrund seiner grossen Bedeutung im Alltag (Wagner, 1982) hier kurz behandelt werden soll (eine umfassende Diskussion findet man in Macho (2010)).



Bsp.3-4: *Simpson-Paradoxon (Todesurteile in Florida zwischen 1972-1979):*

Wir betrachten die Anzahl Todesurteile im Bundesstaat Florida, wobei wir der Frage nachgehen, ob Schwarze für Morde eher zum Tode verurteilt werden als Weisse. Tab. 3-1 zeigt, wie viele Todesstrafen für Weisse und Schwarze ausgesprochen wurden:

Gemäss Tab. 3-1 scheint es keine rassistischen Vorurteile bezüglich der Verhängung der Todesstrafe zu geben, denn von den verurteilten Weissen wurden 3.2% zum Tode verurteilt, während für die Schwarzen der Anteil 2.4% betrug.

Tab. 3-1: *Todesstrafen für Mörder mit weisser und schwarzer Hautfarbe:*

Farbe Mörder	Todesstrafe		Σ	%
	Ja	Nein		
schwarz	59	2448	2507	2.4
weiss	72	2185	2257	3.2
Σ	131	4633	4764	

Wir wollen nun aber auch die Farbe des Opfers mit einbeziehen. Tab. 3-2 zeigt die relevanten Daten.

Tab. 3-2: *Todesstrafen für Mörder mit weisser und schwarzer Hautfarbe unter Berücksichtigung der Hautfarbe des Opfers:*

Farbe Opfer	Farbe Mörder	Todesstrafe		Σ	%
		Ja	Nein		
schwarz	schwarz	11	2209	2220	0.05
	weiss	0	111	111	0.00
weiss	schwarz	48	239	287	16.70
	weiss	72	2074	2146	3.40
Σ		131	4633	4764	

Hier ergibt sich erstaunlicher Weise ein völlig anderes Bild: War das Opfer von schwarzer Hautfarbe, so wurden eher Mörder mit schwarzer als mit weisser Hautfarbe zum Tode verurteilt: 0.05% vs. 0%.

War das Opfer von weisser Hautfarbe, so wurden ebenfalls eher Mörder mit schwarzer als mit weisser Hautfarbe zum Tode verurteilt: 16.7% vs. 3.4%.

Wie kommt es zu dieser Diskrepanz? Diese ergibt sich aus den beiden folgenden Tatsachen:

1. Für einen Mord an einer Schwarzen wird seltener die Todesstrafe ausgesprochen als bei einem Mord an einer Weissen.
2. Schwarze töten eher Schwarze als Weisse. Für Weisse ist es umgekehrt.

Ignoriert man daher die Hautfarbe des Opfers, so scheint es, dass Schwarze milder behandelt würden als Weisse (d.h. es werden insgesamt weniger Todesstrafen gegen sie ausgesprochen).

Dieser Eindruck basiert rein auf der Tatsache, dass Schwarze eher Schwarze ermorden und der Mord an einer Schwarzen weniger oft mit dem Tod bestraft wird.

Schlussfolgerung:

Die Entfernung einer Variable (indem die Werte über die verschiedenen Stufen der Variable summiert werden) kann zu einer völlig falschen Einschätzung führen.

Nach dieser Darstellung der beiden Varianten des HO-Schemas möchte ich zeigen, warum das HO-Schema keine akzeptable Repräsentation der Struktur wissenschaftlicher Erklärungen darstellt.

3.1.2 Einwände gegen das Hempel-Oppenheim-Schema

Im Folgenden soll gezeigt werden, dass die Erfüllung des HO-Schemas weder eine hinreichende noch eine notwendige Bedingung für eine gute wissenschaftliche Erklärung darstellt.

3.1.2.1 DIE ERFÜLLUNG DES HO-SCHEMA IST NICHT HINREICHEND

Wir präsentieren zwei klassische Einwände, welche zeigen, dass das HO-Schema keine hinreichende Bedingung für eine befriedigende Erklärung darstellt.

Das HO-Schema funktioniert auch dann, wenn die Gesetzesaussage für die Erklärung irrelevant ist. Dies führt zu absurden Resultaten.



Bsp.3-5: *HO-Schema und Irrelevanz der Gesetzesaussage:*

Explanandum (E):

Urs Meier wurde letztes Jahr (trotz regelmässigem Sex) nicht schwanger: $\neg S(u)$.

Antezedensbedingungen (A):

1. Urs Meier ist ein Mann: $M(u)$
2. Urs Meier hat regelmässig die Antibabypille seiner Freundin eingenommen: $A(u)$

Gesetzesaussage (G):

Wenn eine Person X ein Mann ist und regelmässig die Antibabypille einnimmt, so wird X nicht schwanger:

$\forall x [M(x) \wedge A(x) \rightarrow \neg S(x)]$

Es ergibt sich die logisch korrekte Schlussfolgerung:

G: $\forall x [M(x) \wedge A(x) \rightarrow \neg S(x)]$

A: $M(u) \wedge A(u)$

E: $\neg S(u)$

Die Erklärung ist insofern unbefriedigend, als die Einnahme der Antibabypille völlig irrelevant für das Nichteintreten einer Schwangerschaft bei Urs Meier ist. Obwohl die Erklärung alle Anforderungen des HO-Schemas an eine wissenschaftliche Erklärung erfüllt, handelt es sich offensichtlich nicht um eine korrekte (wissenschaftliche) Erklärung.

Betrachten wir einen zweiten Einwand, der sich gegen eine grundlegende Annahme Hempels richtet: Hempel nahm an, dass *Vorhersagen* und *Erklärungen* die gleiche logische Struktur besitzen. Der einzige Unterschied besteht demgemäss darin, dass im Falle der *Erklärung* das Explanandum ein vergangenes Ereignis darstellt, während bei *Vorhersagen* das fokale Ereignis in der Zukunft liegt. Aus diesem Grund lässt sich das HO-Schema sowohl auf Erklärungen wie auf Vorhersagen anwenden.

Diese Annahme ist falsch: Vorhersage und Erklärung sind strikt voneinander zu unterscheiden. Bsp.3-6 illustriert das Problem.



Bsp.3-6: *HO-Schema und die Unterscheidung zwischen Erklärung und Vorhersage:*

Betrachten wir folgende Erklärung:

G: Immer wenn die Schwalben tief fliegen, gibt es ein Gewitter.

A: Gestern flogen die Schwalben tief.

E: Gestern gab es ein Gewitter.

Als Vorhersage:

G: Immer wenn die Schwalben tief fliegen, wird es ein Gewitter geben.

A: Die Schwalben fliegen tief.

V: Es wird ein Gewitter geben.

Während die Vorhersage adäquat ist, muss die gegebene Erklärung als völlig unzureichend bezeichnet werden, denn die Tatsache, dass die Schwalben tief fliegen ist keine Erklärung der Tatsache, warum es ein Gewitter gibt. Ähnlich wie in Bsp.3-5 haben wir eine mögliche Erklärung, welche alle Erfordernisse des HO-Schemas erfüllt und dennoch keine adäquate Erklärung darstellt. Dies zeigt, dass die Erfüllung des HO-Schemas nicht hinreicht für eine sinnvolle Erklärung.

Wenden wir uns nun der Frage zu, ob das HO-Schema – wenn schon keine hinreichende – so wenigstens eine notwendige Bedingung für eine gute Erklärung darstellt.

3.1.2.2 DIE ERFÜLLUNG DES HO-SCHEMA IST NICHT NOTWENDIG



Bsp.3-7: Erklärung eines Flugzeugabsturzes nach HO-Schema:

Am 21. Dezember 1988 stürzte ein Verkehrsflugzeug vom Typ Boeing 747-121 der amerikanischen Fluglinie Pan American World Airways über der Ortschaft Lockerbie ab, bei dem alle 259 Insassen ums Leben kamen.

Die Ursache für den Absturz war die Explosion von 340–450g Plastiksprengstoff.

Der findige HO-Theoretiker kann nun die folgende Erklärung vom Typ HO-Schema präsentieren:

- G: Immer wenn ein Plastiksprengstoff von 340-450 g in einem fliegenden Flugzeug vom Typ Boeing 747-212 explodiert, so stürzt das Flugzeug ab.
- A: Ein Plastiksprengstoff von 340-450 g ist in einem fliegenden Flugzeug vom Typ Boeing 747-212 explodiert.
-
- E: Das Flugzeug stürzte ab.

Wir stellen uns nun die folgende Frage:



Frage:

Warum benötigt man für eine Erklärung eine Gesetzesaussage? Man möchte ja nur einen ganz bestimmten Fall erklären und es interessiert vorerst nicht, ob die gegebene Erklärung auch für ähnliche Fälle zutrifft.

Im Hinblick auf Bsp.3-7 stellt sich das Problem wie folgt: Der Absturz wird zufrieden stellend erklärt, wenn man angibt, dass 340-350g Plastiksprengstoff explodierten, die ein Loch in den Rumpf rissen, was weitere (vermittelnde) Ereignisse nach sich zog, die schliesslich zum Absturz führten. Ob dies für alle Flugzeuge bzw. für alle Arten von Plastiksprengstoff zutrifft ist für die Erklärung des betreffenden Falles irrelevant.

Für eine zufrieden stellende Erklärung benötigt man also nicht unbedingt eine Gesetzesaussage. Daher stellt das HO-Schema keine notwendige Bedingung für eine gute Erklärung dar. Dies bedeutet natürlich nicht, dass Wissen um gesetzesmässige Zusammenhänge unwichtig sind für kausale Erklärungen, sondern nur, dass die Erwähnung eines Gesetzes nicht notwendig ist für eine gute Erklärung.

Wir wenden uns nun der Frage zu, was eine zufrieden stellende wissenschaftliche Erklärung ausmacht.

3.1.3 Erklärung als Aufdeckung des kausalen Mechanismus

Worin besteht nun eine zufrieden stellende Erklärung? Die Antwort auf diese Frage findet sich in folgendem Prinzip:



Prinzip 3-1: Erklärung und kausaler Mechanismus:

Die Erklärung eines Phänomens besteht in der Darlegung des relevanten kausalen Mechanismus, der dem Phänomen zugrunde liegt.

Aufgrund von Prinzip 3-1 erkennt man sofort den Grund für die Mängel von Bsp.3-5 und Bsp.3-6: Im ersten Fall werden *kausal irrelevante* Gründe (Einnahme der Antibabypille) in die Erklärung einbezogen. Im zweiten Beispiel werden Gründe (Tiefflug der Schwalben) genannt, welche keine relevanten Ursachen für ein Gewitter darstellen. Um gute Vorhersagen machen zu können, benötigt man nicht notwendigerweise eine genaue Kenntnis der Ursachen. Es genügt, wenn man gute *Prädiktoren* besitzt.



Konzept 3-3: Prädiktor

Als *Prädiktor* bezeichnet man eine Variable, welche zur Vorhersage eines Phänomens dient.

Ein Prädiktor muss keine Ursache des Phänomens darstellen. Ein *nützlicher* Prädiktor muss jedoch einen Zusammenhang mit dem Phänomen aufweisen, indem das Auftreten des Phänomens bei Vorliegen des Prädiktors höher ist als bei dessen Abwesenheit.

Ein guter Prädiktor P_0 zeichnet sich auch dadurch aus, dass er kein Substitut eines anderen (besseren) Prädiktors P_1 darstellt. Letzteres würde bedeuten, dass P_0 über P_1 hinaus die Vorhersage nicht mehr verbessert: Sobald ich P_1 kenne, ist Information über P_0 für die Vorhersage irrelevant. P_1 ist daher ein nützlicherer Prädiktor als P_0 .

Das Verfahren der *multiplen Regression* verwendet eine Reihe von unabhängigen Variablen (UVn) zur Analyse der Varianz einer abhängigen Variable (AV). Die UVn können hierbei entweder einfache Prädiktoren oder aber auch echte Ursachen der AV sein (oder auch eine Mischung aus beiden). Das Verfahren unterscheidet nicht zwischen beiden Arten von UVn (Eine exzellente Darstellung dieser Problematik findet man in Pedhazur, 1982).

Analog besteht eine zufrieden stellende Erklärung des Flugzeugabsturzes in Bsp.3-7 in der Darstellung des relevanten kausalen Mechanismus, der den Absturz herbeiführte. Die Spezifikation dieses Mechanismus macht eine Erwähnung allgemeiner Gesetze überflüssig.

Die Charakterisierung einer zufrieden stellenden Erklärung als Darstellung des relevanten Mechanismus macht deutlich, dass gute wis-

senschaftliche Erklärungen eine enge Beziehung zu den *abduktiven Schlussfolgerungen* aufweisen (Lipton, 2005), denn auch bei letzteren geht es ja um das Finden des korrekten kausalen Mechanismus aus der Menge möglicher Kandidaten (vgl. Konzept 2-1).

Eine Verfechterin des HO-Schemas könnte nun wie folgt argumentieren:

Nehmen wir an, dass die Darstellung des kausalen Mechanismus tatsächlich eine zentrale Komponente einer guten Erklärung darstellt. Aber was spricht dann dagegen, das HO-Schema beizubehalten und zu verlangen, dass die Gesetzhypothese den relevanten Kausalmechanismus enthält.

Dieser Vorschlag ist der Absicht von Hempel, welche zur die Bildung des HO-Schemas führte, diametral entgegengesetzt. Das HO-Schema hat nur dann einen Sinn, wenn man aus der Struktur des Schemas erkennen kann, ob eine formal korrekte Erklärung vorliegt. Die Erfüllung der Forderung, dass es sich bei der Gesetzesaussage um den relevanten kausalen Mechanismus handeln müsse, kann aber nicht aus der Form der Erklärung erschlossen werden. Damit gibt aber die Struktur der Erklärung keinen Aufschluss darüber, ob es sich um eine gute oder schlechte Erklärung handelt. Diese Überlegung stellt aber das HO-Schema als Ganzes in Frage, denn was bringt das Schema, wenn es keinerlei Auskunft darüber gibt, ob die Erklärung gut oder schlecht ist?



Bemerkung: Gesetzesaussagen im HO-Schema:

Diese Problematik, dass das HO-Schema keine Auskunft darüber gibt, ob die Erklärung zufrieden stellend ist, stellt sich auch, wenn man annimmt, dass eine Gesetzesaussage als Prämisse (zusätzlich zur Antezedensbedingung) zu spezifizieren ist. Denn was eine korrekte Gesetzesaussage ausmacht, kann nicht aus der Form der Aussage erschlossen werden.

Es ist meines Wissens bisher noch nicht gelungen, eindeutig zu bestimmen, was eine Gesetzesaussage ausmacht (Curd & Cover, 1998).

Nachdem nun geklärt ist, dass eine gute Erklärung in der Darstellung des relevanten kausalen Mechanismus besteht, wollen wir uns nun verschiedenen Typen von Erklärungen zuwenden, welche in der Psychologie Anwendung finden. Hierbei stellt sich die Frage, ob es sich bei allen diesen Erklärungen um Kausalerklärungen handelt.

3.1.4 Funktionale Erklärungen in der Psychologie

Innerhalb der Psychologie findet man auch so genannte funktionale Erklärungen.



Konzept 3-4: Funktionale Erklärung

Eine *funktionale Erklärung* in der Psychologie besteht in der Angabe der Funktion, welches das zu erklärende Phänomen für das Individuum erfüllt.

Funktionale Erklärungen spielen in der *Evolutionarypsychologie* eine wichtige Rolle. Dieser Zweig der Psychologie erklärt die mentale Ausstattung des Menschen als *funktionales Produkt* der natürlichen Selektion. Demzufolge sind mentale Fertigkeiten als Antworten auf bestimmte adaptive Probleme zu betrachten.



Konzept 3-5: Evolutionspsychologische Erklärung

Eine evolutionspsychologische Erklärung eines psychologischen Mechanismus umfasst die folgenden beiden Aspekte:

1. Eine Spezifikation des vorliegenden adaptiven Problems, welche durch den Mechanismus gelöst werden soll.
2. Eine Explikation der Art und Weise, wie das Problem durch die Ausprägung des psychologischen Mechanismus gelöst wird.

Ein Beispiel mag dies verdeutlichen.



Bsp.3-8: Evolutionspsychologische Erklärung geschlechtsspezifischer Unterschiede im Mechanismus der Partnerwahl:

Gegeben sei der folgende Sachverhalt:

Das Merkmal körperliche Attraktivität hat für die männliche Partnerwahl eine grössere Bedeutung als für die weibliche.

Die Evolutionspsychologie erklärt diese Differenz aufgrund der unterschiedlichen selektiven Rollen von Mann und Frau und der sich heraus ergebenden unterschiedlichen Problemsituationen.

Der Unterschied ergibt sich, indem einerseits die weibliche Investition in den Nachwuchs beträchtlich höher ist als jene durch den Mann (bedingt durch Schwangerschaft und Säugen). Andererseits ist der Mann grundsätzlich in der Lage, eine grössere Anzahl von Nachkommen zu zeugen.

Aufgrund dieser Asymmetrie ergeben sich für die beiden Geschlechter unterschiedliche Zielsetzungen:

- Das Ziel des Mannes besteht darin, möglichst viele Nachkommen zu zeugen (da der Akt der Zeugung relativ wenig Aufwand erfordert). Dadurch erhöht sich die Wahrscheinlichkeit, dass zumindest ein Teil des Nachwuchses überlebt.

- Das Ziel der Frau besteht hingegen darin, den Nachwuchs, in den sie viel investiert hat, möglichst am Leben zu erhalten.

Vereinfacht ausgedrückt lassen sich die beiden Strategien wie folgt zusammenfassen:

Während der Mann die Wahrscheinlichkeit von Nachfolger vorwiegend durch Quantität zu erhöhen anstrebt, wählt die Frau eher eine Strategie der Maximierung durch Qualität der Aufzucht.

Aufgrund dieser unterschiedlichen Strategien bevorzugen Mann und Frau unterschiedliche Qualitäten des Partners:

- Körperliche Attraktivität weist auf Jugend und Gesundheit der Partnerin hin. Folglich maximiert die Paarung mit möglichst vielen attraktiven Partnerinnen die Wahrscheinlichkeit der Fortpflanzung. Der Hinweisreiz körperliche Attraktivität ist daher für den Mann von prominenter Bedeutung.
- Im Hinblick auf die Zielsetzung der Frau sind neben diesem Aspekt noch andere Eigenschaften wichtig. Diese betreffen den Willen und die Fähigkeit des Partners, eine möglichst optimale Aufzucht des Nachwuchses zu gewährleisten. Relevant sind in diesem Zusammenhang (unter anderem) die soziale Stellung sowie die verfügbaren Ressourcen des Mannes. Daher haben diese Aspekte für die Frau einen höheren Stellenwert als für den Mann.

Bei der gegebenen evolutionspsychologischen Erklärung handelt es sich insofern um eine *funktionale Erklärung*, als die beobachtete Asymmetrie hinsichtlich des Stellenwertes der körperlichen Attraktivität des Partners durch die unterschiedliche Funktion, welche dieses Merkmal für die beiden Geschlechter hatte, erklärt wird.

Es stellt sich nun die Frage nach der Beziehung zwischen funktionalen und kausalen Erklärungen. Die Antwort auf diese Frage liefert das folgende Prinzip.



Prinzip 3-2: Kausale und funktionale Erklärungen:

Funktionale Erklärungen lassen sich auf kausale Erklärungen zurückführen. Hierbei wird der kausale Status (oder die kausale Rolle) einer Funktion expliziert.

Funktionale Erklärungen sind daher kein eigenständiger Typ von Erklärungen.

Im Falle der evolutionspsychologischen Erklärungen ist diese Reduktion funktionaler Erklärungen auf kausale offensichtlich. Es handelt sich bei evolutionspsychologischen Erklärungen im Grunde um kausale Erklärungen. Dies sieht man sofort, wenn man noch folgendes Prinzip zur Erklärung hinzufügt:

Den zentralen kausalen Mechanismus zur Herausbildung bestimmter psychologischer Mechanismen bildet die Selektion: Es wurden jene Fertigkeiten und Strategien herausgebildet, welche den zugehörigen Individuen einen selektiven Vorteil verschafften und somit die Ausbreitung der den Mechanismen zugrunde liegenden Gene begünstigten.

Der *relevante kausale Faktor* für einen beobachteten psychologischen Mechanismus ist daher die erhöhte Konkurrenzfähigkeit, welche das Merkmal den Individuen verschafft. Die Explikation dieses kausalen Faktors steht aber im Zentrum der evolutionspsychologischen Erklärung (vgl. Konzept 3-5).



Bemerkung:

Man beachte, dass ein bestimmter psychologischer Mechanismus, welcher zum Zeitpunkt seiner Entstehung wertvoll war und dem Individuum einen Selektionsvorteil verschaffte, unter geänderten Umweltbedingungen seinen Wert verlieren und sogar negative Auswirkungen haben kann.

Als nächstes wenden wir uns der Rolle der so genannten *Hermeneutik* zu, welche oft als Alternative zu kausalen Erklärungen in verschiedenen Sozial- und Geisteswissenschaften betrachtet wird.

3.1.5 Hermeneutik und kausale Erklärungen

Eine Strömung innerhalb der Geistes- und Sozialwissenschaften postuliert die Existenz eines für die Geistes- und Sozialwissenschaften spezifischen Erkenntnisprozesses. So betrachtet Wilhelm Dilthey (1833–1911) das *Verstehen* von Erlebniszusammenhängen als zentral für die Geschichts- und Humanwissenschaften. Hierbei handle es sich um das Verstehen von geistigen Produkten, das sich grundsätzlich von der Erkenntnisgewinnung in den Naturwissenschaften unterscheidet.

Analog stellt Habermas (1968) der instrumentellen Vernunft, wie sie in Technik und Naturwissenschaft vorherrscht, die *Selbstreflexion* gegenüber, welche zur Befreiung von dogmatischen Abhängigkeiten führt. Als paradigmatisch für den Prozess der Selbstreflexion sieht er die psychoanalytische Therapie, in der sich das Individuum durch Verstehen der eigenen Geschichte von Ängsten und Zwängen befreit.



Bsp.3-9: Selbstreflexion und Kausalität:

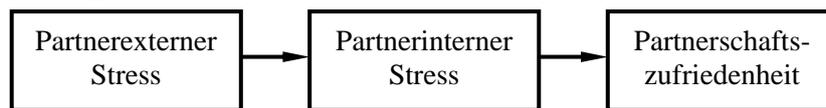
Habermas stellt die psychoanalytische Therapie strikt der somatischen Medizin gegenüber, wobei er wie folgt argumentiert:

Während wir in der technischen Verfügung über Natur, aufgrund unserer Kenntnis der kausalen Zusammenhänge, die Natur für uns arbeiten lassen, trifft die analytische Einsicht die Kausalität des Unbewußten als solche: die Therapie beruht nicht, wie die im engeren Sinne »kausale« der somatischen Medizin, auf einer Indienstnahme der erkannten Kausalzusammenhänge, sie verdankt vielmehr ihre Wirksamkeit der Aufhebung der Kausalzusammenhänge selber (Habermas, 1968, S. 330).

Die Unsinnigkeit dieser Argumentation erkennt man, wenn man sich die Frage stellt, wie es den möglich sein soll, gegebene Kausalzusammenhänge aufzuheben.

Gewöhnlich denkt man hierbei an die Unterbrechung einer Kausalkette, über die eine Wirkung von der Ursachen- auf die Effektvariable vermittelt wird.

Beispiel:



Das dargestellte kausale Modell besagt, dass partnerexterner Stress (z.B. Arbeitsstress) sich auf die Partnerschaftszufriedenheit negativ auswirkt, indem er zu partnerinternem Stress führt. Lässt sich daher die Kette unterbrechen, indem man – z.B. durch ein geeignetes Training der Partner im Umgang mit Stress – die kausale Verbindung zwischen partnerexternem und -internem Stress unterbricht, so verschwindet auch die negative Auswirkung auf die Partnerschaftszufriedenheit.

Die hier dargestellte Vorgehensweise der Unterbrechung einer Kausalkette gehört jedoch eindeutig in den Bereich »der technischen Verfügung über Natur, aufgrund unserer Kenntnis der kausalen Zusammenhänge« und kann somit nicht gemeint sein.

Ausserdem hebt das Training den kausalen Zusammenhang zwischen partnerexternem und -internem Stress nicht grundsätzlich auf. Es bewirkt lediglich, dass die Mechanismen, mittels denen partnerinterner Stress durch partnerexternen Stress erzeugt wird (z.B. durch Ausdruck negativer Emotionen, wie schlechter Laune, Gereiztheit, Apathie, etc.) blockiert werden.

Freud selbst hatte aufgrund seiner »szientistischen Verblendung« (in der er sich gemäss Habermas befand) natürlich niemals in Erwägung gezogen, dass die psychoanalytische Therapie zur Aufhebung von Kausalzusammenhängen dienen könnte. Ihm zufolge basierte die Heilung eines Symptoms auf der Aufhebung der Verdrängung, die er als *kausal notwendig* für das Zustandekommen eines Symptoms betrachtete.

Gemäss der hier vertretenen Ansicht *folgt der Erkenntnisprozess in allen empirischen Wissenschaften der grundsätzlichen gleichen Logik*. Es gibt daher keinen eigenständigen Erkenntnisprozess der Geistes- und Sozialwissenschaften, sofern sie empirisch arbeiten. Im Speziellen handelt es sich bei den Erklärungen in diesen Bereichen ebenfalls um Kausalerklärungen.

Ein Beispiel aus der Geschichtswissenschaft mag dies verdeutlichen.



Bsp.3-10: Historische Erklärung:

Gegeben: Zu erklärender historischer Sachverhalt:

Warum wurde Hitler von Hindenburg zum Reichskanzler ernannt?

Pyta (2007) präsentiert in seiner Hindenburgbiographie eine Reihe von Gründen, welche Hindenburgs anfängliche Bedenken hinsichtlich der Befähigung Hitlers zum Reichskanzler in den Hintergrund drängten:

- Hindenburg hat ein grosses Bedürfnis, sich aus der aktuellen Politik zurückzuziehen, um vorwiegend repräsentative Aufgaben wahrzunehmen. Die ständig wechselnden Präsidialregierungen hatten ihn immer mehr in die aktuelle Tagespolitik verstrickt.
- Hindenburg wollte einen Kanzler, welcher in der Lage wäre, das deutsche Volk zu einigen. Dieser musste natürlich aus dem rechten Lager kommen. Er traute es Hitler zu, diese einigende Funktion zu erfüllen.
- Hitler hatte am ersten Weltkrieg teilgenommen (wenn auch nur als Gefreiter). Er hatte von daher auch den richtigen militärischen »Stallgeruch«.

Geht man davon aus, dass Ziele und Motive, sowie Überlegungen, wie diese Ziele verwirklicht werden können, als Ursachen von Handlungen zu betrachten sind, so können die genannten Gründe als Ursachen der Ernennung Hitlers durch Hindenburg betrachtet werden.

Es gab innerhalb der Philosophie immer wieder Diskussionen, ob man Gründe und Motive als Ursachen für Handlungen betrachten kann. Es

gibt – soweit ich sehe – absolut keine stichhaltigen Gründe, welche dagegen sprechen würden (vergleiche Grünbaum, 1988, Seite 122-146). Nach der hier vertretenen Auffassung besteht eine gute Erklärung eines Phänomens in der Darlegung des kausalen Mechanismus, der zu Erscheinen des Phänomens führte. Hier stellt sich nun eine einfache Frage.



Frage:

Sind alle wissenschaftlichen Erklärungen kausaler Natur?

Die Beantwortung dieser Frage ist schwierig, denn einerseits gibt es in den Naturwissenschaften eindeutig Erklärungen, welche nicht-kausaler Natur sind und die – jedenfalls zum jetzigen Zeitpunkt – nicht auf kausale Erklärungen reduzierbar sind.



Bsp.3-11: Beispiel für eine nicht-kausale Erklärung: Das Pauli-Verbot

Pauli-Verbot, welches besagt, dass innerhalb eines Quantensystems nicht zwei Elemente mit völlig übereinstimmenden Quantenzuständen existieren können. Dieses Prinzip erklärt das Verhalten bestimmter chemischer Elemente, welches auf andere Weise nicht erklärt werden kann (vgl. die gut lesbare Diskussion in Pais, 1993).

Die Frage, ob es eine kausale Erklärung des Pauli-Verbots gibt, kann derzeit nicht beantwortet werden, da eine schlüssige Interpretation der Quantenmechanik aussteht. Wie steht es nun mit Beispielen aus dem Bereich der Psychologie. Betrachten wir hierzu einen möglichen Kandidaten.



Bsp.3-12: Beispiel für eine nicht-kausale Erklärung: Das Phänomen der Regression zur Mitte

Angenommen Sie nehmen im Rahmen Ihres Studiums an einer Prüfung mit multiple-choice Fragen teil. Der Einfachheit halber gebe es pro Frage zwei Antworten, eine falsche und eine korrekte.

Nach absolviertem Test werden die besten und die schlechtesten 10% der Teilnehmer einem zweiten Test mit neuen Fragen aus dem gleichen Fragenpool unterzogen.

Ergebnis:

- Die 10% besten Studentinnen werden im zweiten Test im Durchschnitt ein schlechteres Ergebnis erzielen als im ersten.
- Die 10% schlechtesten Studentinnen werden im zweiten Test im Durchschnitt ein besseres Ergebnis erzielen als im ersten.

Begründung:

Das Ergebnis im ersten Test ist durch zwei Faktoren beeinflusst: (1) Das Wissen der Personen und (2) Zufall (Glück). Das Wissen der Personen wird diesen auch im zweiten Test helfen, aber man kann nicht erwarten, dass ihnen das Glück auch im zweiten Test in diesem Ausmass »hold ist«.

Bemerkung:

Das Phänomen ist am stärksten, falls das Wissen für das Ergebnis überhaupt keine Rolle spielt und das Ergebnis im ersten Test rein durch Zufall entstanden ist.

Die Erklärung in Bsp.3-12 (und des Phänomens der Regression zum Mittel im Allgemeinen) bezieht sich auf die Wirkung von Zufallsfaktoren. Im Allgemeinen gilt es als ein Urteilsfehler für Ereignisse, welche mit Hilfe des Phänomens der Regression zum Mittel erklärt werden können, spezifische Kausalerklärungen zu präsentieren.

Dennoch ist die Erklärung mittels Regression zur Mitte mit der kausalen Natur von Erklärungen völlig vereinbar. Bei dem was oben als Zufall oder Glück bezeichnet wurde handelt es sich klarerweise um kausale Einflüsse. Nur wurden diese nicht systematisch gemessen und – zusätzlich – variieren diese kausalen Einflüsse stark von Situation zu Situation. Der Urteilsfehler besteht darin, systematische situationsinvariante Einflüsse als Erklärung heranzuziehen.

*Bemerkung:*

Urteilsfehler, welche das Phänomen der Regression zum Mittel ignorieren, findet man häufig im Sport.

Die meisten sportlichen Leistungen sind sowohl durch Geschick als auch durch Zufall beeinflusst. Es scheint jedoch so zu sein, dass die meisten Personen (darunter viele Sportexperten) die Rolle der Zufallseinflüsse – d.h. vom kausalen Einflüssen, die stark situationsbedingt sind – unterschätzen.

Angenommen ein Neuling (sei es ein Sportler oder eine Mannschaft) zeigt extrem gute Leistungen, sodass alle Sport-Gazetten aufmerksam werden. Plötzlich jedoch nimmt die Leistung ab. In diesem Fall gibt es dann immer zahlreiche Erklärungen, wie z.B. »der Erfolg ist ihm (ihnen) in den Kopf gestiegen« (Nisbett & Ross, 1980).

Eine einfachere Erklärung besteht darin, dass der Neuling das ihm zugeschriebene Potential nicht besitzt und anfangs durch situationsbedingte Faktoren bevorzugt wurde.

Wir schliessen unsere Untersuchungen von möglichen Erklärungen nicht-kausaler Natur mit dem folgenden Prinzip ab:



Prinzip 3-3: *Wissenschaftliche und kausale Erklärungen:*

Fast jede sinnvolle wissenschaftliche Erklärung ist letztendlich eine kausale Erklärung.

Im nächsten Abschnitt untersuchen wir probabilistische Erklärungen, die in der Psychologie eine wichtige Rolle spielen.

3.1.6 Probabilistische Erklärungen in der Psychologie

3.2 Ebenen der Erklärung eines Informationsverarbeitungssystems

3.3 Psychologische Erklärungen und Reduktionismus

3.4 Scheinerklärungen in der Psychologie

4. Theorien und Theoriendynamik in der Psychologie

Die vorhandenen Methoden haben einen grossen Einfluss auf die Art und Weise, wie Forscherinnen an ihren Forschungsbereich heran gehen. Wie am Beispiel der explorativen Faktorenanalyse dargestellt, können die verwendeten Methoden einen negativen Effekt auf die aktuelle Forschungspraxis ausüben (Dieser Aspekt wird in Kapitel xxx näher behandelt)

4.1 Der Einfluss der Konzeption von Wissenschaft auf die psychologische Forschungspraxis

Eine andere – aber ebenso wichtige – Frage betrifft den Einfluss der Konzeption von Wissenschaft in den Köpfen der Forscher auf deren Forschungspraxis.

4.2 Theorienbildung in der Psychologie

Eine gewisse Rolle für die Entwicklung von Theorien scheinen Gedankenexperimente zu spielen (Kühne, 2005).

Gedankenexperimente

Dunbar (1995)

Gigerenzer (1991)

4.3 Probleme der Prüfung von Theorien in der Psychologie

- Ad-hoc Hypothesen und Alternativerklärungen.
- Systematische Datenauswahl, die Theorien bestätigt.
 - Explorative Studien als konfirmative Studien
 - Überschätzung der Effektstärke durch Unterdrückung von Daten
- Studien, die Theorien nicht bestätigen, werden nicht publiziert, höchstens sie bestätigen einen Teil.

4.3.1.1 KLASSISCHE SIGNIFIKANZTESTS UND DIE PRÜFUNG VON THEORIEN

In den letzten Jahren wurde der klassische Signifikanztest einer umfassenden Kritik unterzogen (Siehe z.B. Harlow, Mulaik, & Steiger 1997; Kline, 2005). Manche Forscher gehen sogar soweit, Signifikanztests jegliche Art von Sinnhaftigkeit abzusprechen (z.B. Schmidt & Hunter, 1997). Ich möchte im Folgenden argumentieren, dass im Zusammenhang mit kritischen Experimenten die Verwendung von Signifikanztests durch aus sinnvoll sein kann (für eine ähnliche Argumentation siehe Mulaik, Raju, & Harshman, 1997).

Bevor ich auf die Anwendung von Signifikanztests im Kontext von kritischen Experimenten eingehe, gebe ich eine kurze Zusammenfassung der zentralen Merkmale von klassischen Signifikanztests.

**Method 4-1: Klassische Signifikanztest**

Klassische Signifikanztests beinhalten zwei wesentliche Konstituenten:

Ein statistisches Modell einer Stichprobenkennwertestatistik (Kennwertestatistik) mit freien Parametern.

Eine statistische Hypothese, welche Aussagen über die freien Parameter des statistischen Modells macht.

Bevor dies weiter ausgeführt wird, einige Erklärungen zum Gesagten:

**Konzept 4-1: Stichprobenkennwerte und Populationsparameter**

3. Deduktive Schlussfolgerungen (Deduktion):



Bsp.4-1: Theorie des assoziative Lernens als wieder belebtes Forschungsprogramm:

5. Pseudowissenschaftliche Praktiken in der Psychologie

6. Verwendete Literatur

- Ackerman, P. L., Beier, M.E., & Boyle, M. O. (2005). Working Memory and Intelligence: The Same or Different Constructs? *Psychological Bulletin*, 131, 30-60.
- Anderson, J. R. (1978). Arguments concerning representations of mental imagery. *Psychological Review*, 85, 249-277.
- Anderson, J. R. (1979). Further arguments concerning representations of mental imagery: A response to Hayes-Roth and Pylyshyn. *Psychological Review*, 86, 395-406.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum. [HPEAD Q-1015]
- Aronson, E., Wilson, T. D. Akert, R. M. (2010) *Social psychology* (7th edition). Boston: Pearson. [HPAED Y-2039A]
- Ein exzellentes Lehrbuch der Sozialpsychologie. Das Buch ist sehr anwendungsorientiert und sehr informativ. Es ist gut und spannend geschrieben. Die Tatsache, dass bereits die 7.Auflage vorliegt zeugt ebenfalls von der Qualität.*
- Balakrishnan, J. D. (1999). Decision processes in discrimination: Fundamental misrepresentations of signal detection theory. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1189-1206.
- Balakrishnan, J. D., & MacDonald, J. A. (2002). Decision criteria do not shift: Reply to Treisman. *Psychonomic Bulletin & Review*, 9, 848-865.
- Balakrishnan, J. D., & MacDonald, J. A. (2010). Decision criteria do not shift: Commentary to Mueller and Weidemann. *Psychonomic Bulletin & Review*, 15, 1022-1030.
- Balzer, W., & Marcou, P. (1989). A reconstruction of Sigmund Freud's early theory of the unconscious. In H. Westmeyer (Ed.), *Psychological theories from a structuralist point of view* (pp. 13-31). Berlin: Springer.
- Baron, J. (2008). *Thinking and deciding*. Cambridge: Cambridge University Press. [FSES M121, TEB-26109, BP2 Ökonomie]
- Ein umfassendes Werk zum Thema Denken und Entscheiden. Leider etwas langatmig und nicht sehr spannende geschrieben.*
- Batchelder, W.H., & Riefer, D.M. (1999). Theoretical and empirical review of multinomial processing tree modeling. *Psychonomic Bulletin & Review*, 6, 57-86.

- Beyerstein, B. (2007). Whence comes the myth that we use 10% of our brains? In T. J. Lawson (Ed.), *Scientific perspectives on pseudoscience and the paranormal: Readings for general psychology* (pp. 47-61). Upper Saddle River, NJ: Prentice. [HPAED Q-1951]
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bower, G., & Trabasso, T. (1963). Reversals prior to solutions in concept identifications. *Journal of Experimental Psychology*, *60*, 409-418.
- Clark, S. E., & Gronlund, S. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, *3*, 37-60.
- Clogg, C. C. (1995). Latent class models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 311-359). New York: Plenum Press. [HPAED PC-142]
- Cowles, M. P. (2001). *Statistics in psychology: A historical perspective*. (2nd edition) Mahwah, NJ: Erlbaum. [HPEAD PC-322]
- Das Buch von Cowles erzählt die Geschichte der Entwicklung der zentralen statistischen Verfahren der Psychologie, sowie ihrer wichtigsten Proponenten, wie Karl Pearson, Ronald Fischer und Jerzy Neyman. Das Buch ist stellenweise amüsant jedoch teilweise auch beschwerlich zu lesen. Letzteres liegt vor allem daran, dass Cowles die verwendeten Symbole nicht immer erklärt.*
- Curd, M., & Cover, J. A. (1998). *Philosophy of science: The central issues*. New York: Norton. [DOKPE Q-175; TSB-8961]
- Eine Sammlung der wichtigsten Aufsätze zur Philosophie der Wissenschaft. Enthält viele Klassiker des Gebiets mit Erklärungen der Autoren. Hier werden alle wichtigen Probleme der Wissenschaftstheorie behandelt.*
- Dawes, R. M. (1988/2001). *Rational choice in an uncertain world*. Philadelphia: Harcourt Brace Jovanovich. [HPAED Q-1160, HPAED Q-1819]
- Ein interessantes Buch, welches ausgewählte Themen zu Entscheiden und Urteilen darstellt. Gut und kompetent geschrieben, informativ.*
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press. [HPAED Q-1510]

Ein sehr interessantes Buch zum Thema Expertentum in der Psychologie. Die hier ausgesprochenen Thesen werden vielen klinischen Psychologinnen nicht gefallen!

DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: Theoretical developments with applications to recognition memory. *Psychological Review*, 109, 710-721.

DeCarlo, L. T. (2007). The mirror effect and mixture signal detection theory. *Journal of Experimental Psychology, Learning, Memory, & Cognition*, 33, 18-33.

Dehaene, S. (2010). Lesen: Die grösste Erfindung der Menschheit und was dabei in unseren Köpfen passiert. München: Knaus.

Das Buch bietet einen guten Überblick über psychologische, neuropsychologische und evolutionspsychologische Erklärungen des Lesens. Es ist einfach und interessant geschrieben.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories In R. J. Sternberg, & J. E. Davidson (Eds.), *The nature of insight* (356-395). Cambridge, Mass.: MIT-Press. [HPAED Q-1427]

Ekman, P., & Friesen, W. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists Press. [HPAED Pb-173].

Epstein, S. (1994). Trait theory as personality theory: Can a part be as great as the whole? *Psychological Inquiry*, 5, 120-122.

Ernst, E., & Singh, S. (2009). *Gesund ohne Pillen - was kann die Alternativmedizin?* München: Hanser.

Eine informative, geistreiche und unterhaltsame Darstellung der Geschichte und Hintergründe von verschiedenen alternativmedizinischen Praktiken, wie Akupunktur, Homöopathie oder Chiropraxis.

Estes, W. K. (1960). Learning theory and the new "mental chemistry". *Psychological Review*, 67, 207-233.

Feyerabend, P. K. (1995). *Wider den Methodenzwang* (5. Auflage). Frankfurt: Suhrkamp Taschenbuch Wissenschaft 597. [SPAED B-471]

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, Mass.: MIT Press.

Gallistel, C. R., & King, A. P. (2009). *Memory and the computational brain: Why cognitive science will transform neuroscience*. Chichester: Wiley-Blackwell.

- Garcia, J., McGowan, B. K., Ervin, F., & Koelling, R. (1968). Cues: Their relative effectiveness as reinforcers. *Science*, *160*, 794-795.
- Gigerenzer, G. (1991). From tools to theory: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*, 254-267.
- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press. [SPAED L-2-508]
- Eine Sammlung von Aufsätzen zu verschiedensten Themen über Urteile und Urteilsfehler. Die meisten der Aufsätze sind bereits in Journalen publiziert worden. Dieses Buch ist nicht dazu gedacht als Ganzes gelesen zu werden. Es handelt sich um eine Fortsetzung des Werkes von Kahneman, Slovic und Tversky (1982) [siehe unten].*
- Gopnik, A., Glymour, C., Sobel, D. M., Schultz, L. E., & Kushnir, T. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3-32.
- Grünbaum, A. (1988). *Die Grundlagen der Psychoanalyse: Eine philosophische Kritik*. Stuttgart: Reclam.
- Habermas, J. (1968). *Erkenntnis und Interesse*. Frankfurt am Main: Suhrkamp.
- Harlow, L. L., Mulaik, S. A., & Steiger, J. L. (1997). *What if there were no significance tests?* Mahwah, NJ: Psychology Press. [HPAED PC-284]
- Hilford, A., Glanzer, M., Kim, K., & DeCarlo, L. T. (2002). Regularities of source recognition: ROC analysis. *Journal of Experimental Psychology: General*, *131*, 494-510.
- Hübner, K. (1985). *Die Wahrheit des Mythos*. München: C.H. Beck.
- Huber, O., Beutter, C., Montoya, J., & Huber, O. W. (2001). Risk-defusing behaviour: Towards an understanding of risky decision making. *European Journal of Cognitive Psychology*, *13*, 409-426.
- Huber, O., Wider, R., & Huber, O. W. (1997). Active information search and complete information presentation in naturalistic risky decision tasks. *Acta Psychologica*, *95*, 15-29.
- Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press. [HPAED Q-508]
- Eine Sammlung von klassischen Aufsätzen zu verschiedensten Themen über Urteile und Urteilsfehler. Informativ, aber teilweise etwas langatmig.*

- Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp.279-296). New York: Appleton-Century-Croft.
- Kane, M. J., Hambrick, D. Z., Tubolski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visospatial memory span and reasoning.. *Journal of Experimental Psychology: General*, 133, 189-217.
- Kelley, H. H. (1972). Causal schemata and the attribution process. In E.E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (1972), *Attribution: Perceiving the causes of behavior* (pp.151-174). Morristown, NJ: General Learning Corporation.
- Kitcher, P. (2009). *Mit Darwin leben: Evolution, Intelligent Design und die Zukunft des Glaubens*. Frankfurt: Suhrkamp.
- Kline, R. B. (2005). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association. [HPAED PC-287]
- Ein exzellentes Buch über Alternativen zu klassischen Signifikanztests. Einfach geschrieben und sehr informativ.*
- Kosslyn, S. M., Pinker, S., Smith, G. E., & Shwartz, S. P. (1979). On the demystification of mental imagery. *Behavioral and Brain Sciences*, 2, 535-548.
- Krampen, G (1991). *Fragebogen zu Kompetenz- und Kontrollüberzeugungen (FKK)*. Göttingen: Hogrefe.
- Kuhn, T. S. (1976). *Die Struktur wissenschaftlicher Revolutionen*. Frankfurt a. M.: Suhrkamp.
- Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge, UK: Cambridge University Press.
- Lakatos, I., & Musgrave, A. (1974). *Kritik und Erkenntnisfortschritt*. Braunschweig: Vieweg.
- Lawson, T. J. (2007). *Scientific perspectives on pseudoscience and the paranormal: Readings for general psychology*. Upper Saddle River, NJ: Prentice. [HPAED Q-1951]
- Eine Sammlung von Aufsätzen welche verschiedene Ursachen für den Glauben an pseudowissenschaftliche Praktiken aufzeigt. Die Einführung präsentiert wesentliche Charakteristika pseudowissenschaftlicher Praxis.*

- Lilienfeld, S. O., Lynn, S. J., Lohr, J. M. (2003). *Science and pseudoscience in clinical psychology*. New York: Guilford Press. [HPAED X-1615].
- Ein Herausgeberbuch, in dem verschiedene Spezialisten Probleme, Kontroversen und auch zwielichtige Praktiken in der Psychotherapie behandeln.*
- Lilienfeld, S. O., Lynn, S. J., Rusco, J., & Beyerstein, B. L. (2010). *Fifty great myth of popular psychology: Shattering widespread misconceptions about human behavior*. Chichester: Wiley-Blackwell, [HAED Q-1927]
- Eine informative und über weite Strecken sehr amüsante Darstellung verschiedener populärer fehlerhafte psychologischer Theorien.*
- Lipton, P. (2005). *Inference to the best explanation* (2nd edition). London: Routledge. [HPAED PA-]
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass.: Addison Wesley. [HPEAD Pb-72]
- Macho, S. (2007). Feature sampling in detection: Implications for the measurement of perceptual independence. *Journal of Experimental Psychology: General*, 136, 133-153.
- Macho, S. (2010). *Urteilsfehler*. Vorlesungsmanuskript Universität Fribourg. Verfügbar unter:
<http://www.unifr.ch/psycho/site/assets/files/Allg/Macho/Manuskripte/Urteilsfehler.pdf>
- Massaro, D. W. & Friedman, D. (1991). Adaptive rationality and identifiability of psychological processes. *Behavioral and Brain Sciences*, 14, 499-500.
- McCrae, R. R., & John, O.P. (1990). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215. Available from:
<http://00t0holtgrav.iweb.bsu.edu/623/ffmarticle.pdf>
- Mueller, S. T., & Weidemann, C. T. (2008). Decision noise: An explanation for observed violations of signal detection theory. *Psychonomic Bulletin & Review*, 15, 465-494.
- Mulaik, S. A., Raju, N. S., & Harshman, R. A. (1997). There is a time and a place for significance testing. In L.L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (Chapter 4, pp. 65-115). Mahwah, NJ: Psychology Press.

Neapolitan, R. E. (1990). *Probabilistic reasoning in expert systems: Theory and algorithms*. New York: Wiley. [DIUF 1.8.2, TIB-5107]

Eine exzellente Darstellung von probabilistischem Schlussfolgern in Expertensystemen. Erfordert solide Grundkenntnisse in Wahrscheinlichkeitstheorie.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106, 226-254.

Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of human judgment*. Englewood Cliffs, NJ: Prentice-Hall. [HPAED Q-1060]

Ein Klassiker und eines meiner Tophits! Trotz des hohen Alters des Buches immer noch sehr informativ. Sehr gut geschrieben und unterhaltsam. Noch immer sehr empfehlenswert.

Nisbett, R. E. & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.

Pais, A. (1993). *Niels Bohr's times: In physics, philosophy and polity*. Oxford: Clarendon Press.

Patry, P. (2008). *Wissenschaft und Pseudowissenschaft: Ein Beitrag zur Abgrenzungsproblematik*. Saarbrücken: Verlag Dr. Müller,

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd edition). New York: Holt, Rinehard & Winston. [HPAED PC-10A]

Ein exzellentes Lehrbuch zu multipler Regression und anderen damit in Zusammenhang stehenden Verfahren. Trotz des hohen Alters des Buches immer noch sehr lesenswert.

Plous, S. (1993). *The psychology of judgment and decision making*. New York: McGraw-Hill. [HPAED Q-1157]

Eine kompakte und sehr gut geschriebene Zusammenfassung der wichtigsten Aspekte von Entscheidungen und Urteilen. Informativ und unterhaltsam. Das Buch hat den William James Book Award erhalten.

Pohl, R. (2004). *Cognitive Illusions: A handbook of fallacies and biases in thinking, judgment and memory*. Hove, UK: Psychology Press. [HPAED Q-1713]

Gibt einen Überblick zur Psychologie der Urteilsfehler. Das Besondere and dem Buch besteht darin, dass zu den besprochen Themen konkrete Anweisung zur Durchführung einer Studie gegeben werden. Ein

Nachteil besteht in der sehr schlecht geführten Rationalitätsdebatte: Das Buch will es jedem Recht machen. Leider enthält das Buch auch Fehler.

Pylyshyn, Z.W. (1979a). Imagery theory: Not mysterious - just wrong. *Behavioral and Brain Sciences*, 2, 561-563.

Pylyshyn, Z. W. (1979b). Validating computational models: A critique of Anderson's indeterminacy of representation claim. *Psychological Review*, 86, 383-394.

Pylyshyn, Z.W. (2002). Mental imagery? In search of a theory. *Behavioral and Brain Sciences*, 25, 157-238.

Pyta, W (2007): *Hindenburg. Herrschaft zwischen Hohenzollern und Hitler*. Berlin: Siedler.

Quine, W. V. O. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60, 20-43. Available from: <http://www.ditext.com/quine/quine.html>.

Radvansky, G. A. (1999). Memory retrieval and suppression: The inhibition of situation models. *Journal of Experimental Psychology: General*, 128, 563-579.

Reichenbach. H. (1938). *Experience and prediction. An analysis of the foundations and the structure of knowledge*. Chicago: University of Chicago Press.

Rescorla, R. A., & Wagner, A. R. (1972). A theory in Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In: A. H. Blach & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Croft.

Rock, I. (1957). The role of repetition in associative learning. *The American Journal of Psychology*, 70, 186-193.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structures of categories. *Cognitive Psychology*, 7, 573-605.

Rosenzweig, P. (2007). *The halo effect and eight other business delusions that deceive managers*. New York. Free Press. [HPAED W-2077]

Das Buch behandelt die negativen Auswirkungen des Halo-Effekts auf (halb-) wissenschaftliche Untersuchungen und journalistische Berichte über die Ursachen des Erfolgs von Unternehmen. Neben der Wirkung des Haloeffekts werden noch andere Probleme und Fehler in der Beurteilung von Manager und Unternehmen aufgezeigt.

Das Buch ist unterhaltsam und leicht verständlich geschrieben und es ist äusserst interessant und lehrreich.

Ruben, D.-H. (1993). *Explanation*. Oxford: Oxford University Press. [HPEAD PA-76]

Salmon, W. C. (1989). Four decades of scientific explanation. In P. Kitcher, & W. C. Salmon (Eds.), *Scientific explanation: Minnesota studies in the philosophy of science* (Vol. 13, pp. 3-219). Minneapolis: University of Minneapolis Press.

Salmon, W. C. (2006). *Four decades of scientific explanation*. Pittsburg: Pittsburg University Press. [HPAED P-96]

Schmidt, F. L., & Hunter, J. E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In L.L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (Chapter 3, pp. 37-64). Mahwah, NJ: Psychology Press.

Simpson, E. H. (1951): The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B*, 13. 238–241.

Stanovich, K. E. (2010). *Decision making and rationality in the modern world*. Oxford, UK: Oxford University Press. [HPAED Q-1920]

Kurz, informativ und gut geschrieben. Enthält auch eine gute Diskussion zur Rationalitätsdebatte. Empfehlenswert.

Stegmüller, W. (1985). *Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie II: Theorie und Erfahrung* (2. korrigierte Auflage). Berlin: Springer.

Suppe, F. (1977). *The structure of scientific theories* (2nd edition). Urbana: University of Illinois Press. [HPAED PA-70]

Susskind, L. (2010). *Der Krieg um das Schwarze Loch: Wie ich mit Stephen Hawking um die Rettung der Quantenmechanik rang*. Berlin: Suhrkamp.

Sutherland, S. (2007). *Irrationality*. London: Pinter & Martin. [HPAED Q-1900]

Amüsant und gut geschrieben. Gibt einen guten Überblick über eine sehr breite Palette von Urteilsfehler. Einfach zu lesen.

Thaler, R., & Sunstein, C. R. (2009). *Nudge: Wie man kluge Entscheidungen anstösst*. Berlin: Econ. [Kantonsbibliothek Fribourg]

- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34, 273-286.
- Treisman, M. (2002). Is signal detection theory fundamentally flawed? A response to Balakrishnan (1998a, 1998b, 1999). *Psychonomic Bulletin & Review*, 9, 845-857.
- Van Zandt, T. (2000). ROC Curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26, 582-600.
- Wagner, C. L. (1982). Simpson's Paradox in real life. *The American Statistician*, 36, 46-48.
- Weidemann, C. T. & Mueller, S. T. (2008). Decision noise may mask criterion shifts: Reply to Balakrishnan and MacDonald (2008). *Psychonomic Bulletin & Review*, 15, 1031-1034.
- Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, 121, 222-236.
- Watkins, J. W. N. (1984). *Science and scepticism*. Princeton, NJ: Princeton University Press.
- Für Anfänger zu schwierig. Benötigt Voraussetzungen in Logik und Wahrscheinlichkeitstheorie.*
- Westmeyer, H. (1989). *Psychological theories from a structuralist point of view*. Berlin: Springer. [HPAED P-71]
- Westmeyer, H. (1992). *The structuralist program in psychology: Foundations and application*. Seattle: Hogrefe. [HPAED PA-37]
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik: Ein Lehrbuch zur Psychologischen Methodenlehre*. Göttingen: Hogrefe. [HPAED PA-56, PA-56+]
- Eine einfache und gut geschriebene Einführung in die Wissenschaftstheorie für Psychologen.*
- Westfall, R. S. (1980). *Never at rest: A biography of Isaac Newton*. Cambridge: Cambridge University Press.
- Die Newton-Biographie.*

