

Denken mit und über Kausalmodelle

Dissertation

zur Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg - August - Universität zu Göttingen

Vorgelegt von

York Christoph Hagmayer
aus Hilden

Göttingen 2000

D 7

Referent: Prof. Dr. M. R. Waldmann

Korreferent: Prof. Dr. G. Lür

Tag der mündlichen Prüfung: 01.02.2001

3.	EMPIRISCHE UNTERSUCHUNGEN	_	_	_	_	_	93
3.1	Experiment zur Bestimmung von Kausalzusammenhängen	_	_				94
	Experiment 1						94
3.2	Experimente zur Prüfung kausaler Strukturhypothesen	_	_				104
3.2.1	Experiment 2						107
3.2.2	Experiment 3						115
3.2.3	Experiment 4						121
3.2.4	Experiment 5						126
3.2.5	Befragungsstudie						141
3.3	Experimente zur Vorhersage strukturell implizierter Zusammenhänge	_					146
3.3.1	Experiment 6						148
3.3.2	Experiment 7						153
3.3.3	Experiment 8						158
4.	DISKUSSION	_	_	_	_	_	167
4.1	Zusammenfassung der Befunde	_	_	_	_	_	168
4.1.1	Implizite vs. explizite Sensitivität für strukturelle Implikationen						171
4.1.2	Denken mit und über Kausalmodelle						172
4.1.3	Analogie zwischen physikalischem und kausalem Denken						173
4.2	Evolutionenpsychologische Sichtweise kausalen Denkens	_	_				175
4.2.1	Lernen durch Beobachten						176
4.2.2	Lernen durch Handeln						182
4.2.3	Anforderungen der Wissenschaft an kausales Denken						184
4.2.4	Denken mit und über Kausalmodelle – Reprise						186
	LITERATURVERZEICHNIS	_	_	_	_	_	188

DENKEN MIT KAUSALMODELLEN – EIN AKTUELLES BEISPIEL

Am 25. Juli 2000 stürzte in Gonesse bei Paris eine Concorde kurz nach dem Start ab. Bei der Explosion starben alle 113 Passagiere und Besatzungsmitglieder. Videoaufnahmen des Starts zeigten, wie das Flugzeug bereits beim Abheben einen langen Feuerschweif hinter sich herzog.

Am 31. August 2000 legte die Untersuchungskommission einen vorläufigen Bericht über die Unfallursache vor. In diesem wird ein Kausalmodell des Absturzes entwickelt. Die untenstehende Abbildung zeigt eine vereinfachte Darstellung des Unfallhergangs.

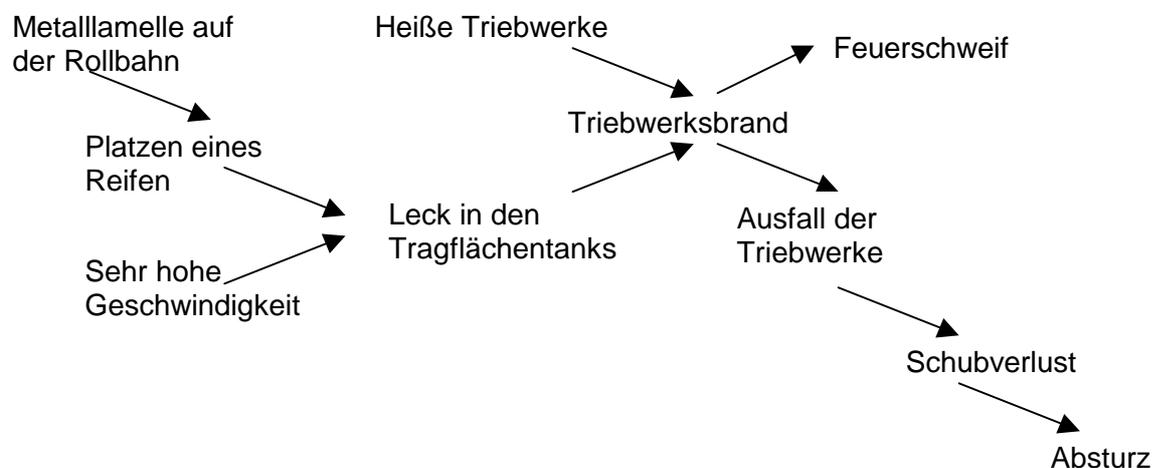


Abb. 1.1: Vereinfachtes Kausalmodell des Concorde-Absturzes bei Paris

Aufgrund einer Metalllamelle auf der Startbahn platzte mindestens einer der Reifen des Fahrwerkes der Concorde. Die Bruchstücke schlugen aufgrund der hohen Startgeschwindigkeit Lecks in die Tragflächentanks. Das austretende Kerosin entzündete sich, wodurch die weithin sichtbaren Flammen entstanden. Eines der Triebwerke wurde daraufhin abgeschaltet, das andere fiel kurz danach aus. Durch den Schubabfall konnte die eingeleitete Linkskurve für eine Notlandung nicht ausgeführt werden. Das Flugzeug stürzte ab und explodierte aufgrund der großen Kerosinmenge in einem enormen Feuerball. Für Details siehe den Bericht unter www.bea-fr.org.

0. EINFÜHRUNG

Der Untersuchungsbericht zum Concorde-Absturz liefert ein aktuelles Beispiel für das Denken mit Kausalmodellen. Kausalmodelle wie dieses werden auch im Alltag ständig von uns entwickelt, erweitert, verändert und manchmal auch verworfen. Wie funktioniert der neue Videorecorder? Warum hat B. seine Frau verlassen? Wie konnte der Rechtsextremismus in den neuen Ländern so stark werden? Kausalmodelle liefern vorläufige Antworten auf diese Fragen. Sie helfen uns, das Geschehen um uns herum zu verstehen und erfolgreich in unserer Umwelt zu agieren.

In den letzten zehn Jahren wurden von zwei Seiten verstärkte Anstrengungen unternommen, Kausalmodelle näher zu untersuchen. Computerwissenschaftler waren seit langem auf der Suche nach effektiven Repräsentationen von Kausalzusammenhängen und erfolgreichen Schlussalgorithmen. Dabei sind sie auf Bayesianische Netzwerkmodelle gestoßen (Pearl, 1988). Seither ist die Entwicklung förmlich explodiert (Glymour & Cooper, 1999; Spirtes et al., 1993) und hat mit der Entwicklung eines eigenen kausalen Kalkulus, der die klassische Wahrscheinlichkeitsrechnung erweitert, einen vorläufigen Höhepunkt gefunden (Pearl, 2000). Kausalmodelle im Sinne dieser Computermodelle sind universell anwendbar, können über Lernen verändert werden und lassen sich statistisch überprüfen. Insbesondere erlauben sie es, Aussagen darüber zu machen, unter welchen Bedingungen aus Beobachtungs- oder experimentellen Daten valide Schlüsse auf Kausalzusammenhänge gezogen werden können.

Eine zweite Forschungslinie zu Kausalmodellen kommt aus der Psychologie (Cheng & Novick, 1990, 1992; Kelley, 1973; Waldmann, 1994, 1996; Waldmann & Holyoak, 1992). Diese Arbeiten orientieren sich an den Erkenntnissen aus der Philosophie (Cartwright, 1989; Eells, 1991; Hume, 1987; Mackie, 1974; Mill, 1968; Salmon, 1980). Sie beschäftigen sich mit der Frage, inwieweit unsere Urteile in Bezug auf Kausalität den normativ richtigen entsprechen. Dabei dienen häufig die von der Wissenschaftstheorie und den Computerwissenschaften aufgestellten Normen als Vergleichsmaßstab. In den letzten Jahren haben sich diese beiden Forschungstraditionen immer stärker miteinander vernetzt. So wurden z.B. Bayesianische Modelle kausalen Denkens in der Psychologie entwickelt (Waldmann & Martignon, 1998) und neue Bestimmungsmaße für Kausalität aus der Psychologie in die Computerwissenschaften übernommen (Cheng, 1997; Pearl, 1999).

Zielsetzung der vorgelegten Arbeit ist es, die normativ-formale wie die psychologische Seite von Kausalmodellen vorzustellen. Dabei wird es um zwei Hauptfragen gehen:

„Worin unterscheiden sich verschiedene Kausalmodelle?“ (normative Fragestellung) und

„Sind Menschen für die Unterschiede zwischen den Modellen sensitiv?“ (psychologische Fragestellung).

Das erste Kapitel beschäftigt sich mit der ersten der beiden Hauptfragen. Qualitative, formale und quantitative Aspekte von Kausalmodellen werden besprochen und Unterschiede zwischen drei grundlegenden Modellen herausgearbeitet. Im Einzelnen wird es dabei um Fragen gehen wie: Was sind Kausalmodelle? Welche Funktionen haben diese? Worin unterscheiden sich verschiedene Modelle? Wie lassen sich diese formalisieren? Wie kann die Stärke eines Kausalzusammenhangs quantifiziert werden?

Kapitel 2 wendet sich der zweiten der Frage zu, ob Personen für die Unterschiede zwischen verschiedenen Kausalmodellen sensitiv sind. Diese Frage lässt sich nicht per se beantworten, sondern nur in Bezug auf konkrete Aufgabenstellungen. Wenn Personen sensitiv für die Unterschiede zwischen verschiedenen Kausalmodellen sind, dann sollte sich dies unter anderem zeigen beim Lernen von Kausalzusammenhängen und Kausalmodellen, bei Diagnosen wie Prognosen, bei der Überprüfung und der Generierung von Kausalmodellen. Für jeden dieser Fälle wird zunächst das normativ richtige Vorgehen besprochen und danach werden, sofern vorhanden, zugehörige Befunde aus der Forschung vorgestellt. Aus den Fragen, die im Anschluss an die vorhandene Literatur noch offen bleiben, werden die Untersuchungen abgeleitet, die im Rahmen dieser Arbeit durchgeführt wurden.

Kapitel 3 stellt diese Studien im einzelnen vor. Alle Untersuchungen überprüfen auf unterschiedliche Weise, ob Probanden sensitiv für die Unterschiede zwischen verschiedenen Kausalmodellen sind. Das erste Experiment beschäftigt sich mit der Frage, ob Probanden bei der Bestimmung der Stärke eines Kausalzusammenhangs dessen Einbettung in ein Kausalmodell berücksichtigen. Normativ gesehen sind die umgebenden Kausalmodelle entscheidend dafür, welche statistischen Zusammenhänge indikativ für Kausalität sind. Die Experimente 2 bis 5 untersuchen, ob Probanden die Fähigkeit haben, Kausalmodelle anhand von Daten zu überprüfen. Hierfür gibt es neben einer kleinen Anzahl korrekter Vorgehensweisen auch eine Vielzahl denkbarer Heuristiken. In den Experimenten 6, 7 und 8 geht es um die Frage, ob Probanden in der Lage sind, separat voneinander gelernte Kausalzusammenhänge zu verschiedenen Modellen zu verbinden und aus diesen Modellen Schlussfolgerungen über die Zusammenhänge zwischen den Ereignissen abzuleiten.

Kapitel 4 versucht aus der Vielzahl der vorliegenden Befunde ein einheitliches Bild zu zeichnen. Dabei wird die im Titel dieser Arbeit angedeutete Unterscheidung zwischen dem Denken mit und dem Denken über Kausalmodelle ein wichtiges Ordnungskriterium sein. Eine Analogie zum physikalischen Denken soll das Gesagte weiter veranschaulichen. Zum Abschluss dieser Arbeit wird eine evolutionspsychologische Sichtweise kausalen Denkens vorgestellt. Dabei wird kausales Denken als Adaptation an die Anforderungen unserer natürlichen Umwelt betrachtet werden. Spekulationen über einen möglichen evolutionären Hintergrund der aufgezeigten Unterschiede zwischen einem Denken mit und einem Denken über Kausalmodelle werden diese Arbeit beschließen.

Am Ende dieser Arbeit darf kein endgültiges Urteil über Kausalmodelle und ihren Einfluss auf unser Denken erwartet werden. Noch gibt es zu viele lose Enden, die es zu verknüpfen gilt, bevor eine Theorie entwickelt werden kann, welche die gesamte Vielfalt des Denkens mit und über Kausalmodelle umfasst.

1. GRUNDLAGEN

Dieses erste Kapitel gibt einen Überblick über qualitative, formale und quantitative Aspekte von Kausalmodellen. Dabei werden die Unterschiede zwischen drei grundlegenden Kausalmodellen herausgearbeitet. Dieses Kapitel geht nicht auf psychologische Aspekte ein. Das wird die Aufgabe der folgenden Kapitel sein. Die unten stehende Abbildung zeigt die Struktur des ersten Kapitels im Einzelnen.

1. GRUNDLAGEN		
1.1 Qualitative Aspekte von Kausalmodellen	1.2 Formale Aspekte von Kausalmodellen	1.3 Quantitative Aspekte von Kausalmodellen
	1.2.1 Lineare Strukturgleichungen	1.3.1 Bestimmung der Stärke einzelner Kausalzusammenhänge
	1.2.2 Bayesianische Modelle	1.3.2 Bestimmung von Kausalzusammenhängen innerhalb eines Kausalmodells
		1.3.3 Berechnung von Kausalmodellen
		1.3.4 Bestimmung struktureller Implikationen
1.4 Zusammenfassung		

Abb. 1.2: Struktur Kapitel 1

1.1 Qualitative Aspekte von Kausalmodellen

Ziel dieses Abschnittes ist es, Kausalmodelle einzuführen. Dabei werden anhand einer Reihe von einfachen Fragen wesentliche Eigenschaften von Kausalmodellen vorgestellt. Eine Vielzahl von Beispielen soll das Gesagte veranschaulichen.

Was sind Kausalmodelle?

Kausalmodelle sind Repräsentationen von bzw. Annahmen über Ursache-Wirkungszusammenhänge in der Welt. Sie werden dabei in der Regel in der Form von Graphen dargestellt. Das Kausalmodell des Concordeabsturzes ist ein Beispiel hierfür. In diesen Graphen werden nur die für einen Sachverhalt kausal relevanten Ereignisse

aufgenommen. Stets vorhandene Randbedingungen werden vernachlässigt, auch wenn diese notwendig sind. So wird in das Kausalmodell des Absturzes weder das Vorliegen von Sauerstoff als eine notwendige Bedingung für Feuer, noch die Gravitation als Bedingung für einen Absturz mit aufgenommen. Die Pfeile innerhalb des Graphen stehen für direkte Kausalbeziehungen. Beispielsweise verursacht der Triebwerksbrand unmittelbar den Feuerschweif. Nicht aufgenommen werden in ein Kausalmodell mittelbare Kausalzusammenhänge, z.B. dass der geplatzte Reifen zum Absturz führte. Da die beiden Ereignisse über mehrere andere Kausalbeziehungen miteinander verbunden sind, braucht der daraus resultierende Zusammenhang nicht mehr gesondert vermerkt werden. Auch werden in das Modell keine rein statistischen Beziehungen aufgenommen. Die Auswirkungen des Triebwerksbrandes korrelieren natürlich alle miteinander. Da dies aber die Folge der Kausalzusammenhänge ist, muss dies nicht separat repräsentiert werden. Diese Beschränkung auf die kausal relevanten Ereignisse und ihre unmittelbaren Kausalrelationen ist ein wesentliches Merkmal von Kausalmodellen. Dies erlaubt eine sehr sparsame Repräsentation eines Sachverhaltes. Außerdem lassen sich alle zu erwartenden mittelbaren Kausalzusammenhänge und statistische Beziehungen über die beteiligten Kausalzusammenhänge rekonstruieren. Wie dies möglich ist, werden wir in später genauer kennen lernen.

Kausalmodelle haben den Status von Theorien bzw. Hypothesen. Sie stellen dar, wie die kausale Struktur der Welt sein könnte. Als Theorien können sie durch entsprechende empirische Belege bestätigt, aber nie endgültig verifiziert werden. Zum Beispiel weist der Bericht zum Concorde-Absturz ausdrücklich darauf hin, dass das aufgezeigte Szenario nur ein möglicher Unfallhergang ist. Wie der Abschlußbericht der Untersuchungskommission des Absturzes des TWA Jumbos vor Rhode Island gezeigt hat, ist bei Einzelfällen häufig auch nach jahrelanger Forschung eine völlige Sicherheit nicht zu erreichen.

Kausalmodelle repräsentieren allgemeine Ursache-Wirkungsbeziehungen. Sie beziehen sich nicht auf Einzelfälle. So gilt z.B. für das Modell des Concorde-Absturzes, dass Gegenstände auf der Rollbahn generell mit einer bestimmten Wahrscheinlichkeit zum Platzen eines Reifens führen. Die Bruchstücke wiederum führen ganz allgemein bei hoher Geschwindigkeit mit einer gewissen Wahrscheinlichkeit zu einem Leck in der Tragfläche usw.. Das Modell repräsentiert eine potentielle Ursache-Wirkungskette, die für alle Flugzeugstarts jeden Tag gültig ist. Glücklicherweise lief die Kette bisher nur im Falle des Concorde-Absturzes Ende Juli von Anfang bis Ende ab. Die beteiligten Kausalzusammenhänge sind extrem schwach. So kommt es praktisch nie vor, dass ein geplatzter Reifen zu einem Leck in der Tragfläche führt. Von dem her war es und ist es extrem unwahrscheinlich, dass die Kette bis zum Absturz durchlaufen wird. Dennoch war

dieses Kausalmodell Grund genug für die Luftfahrtbehörden, weitere Starts der Concorde zu untersagen.

Welche Funktionen haben Kausalmodelle?

Kausalmodelle helfen uns Phänomene zu verstehen. Ein Sachverhalt wird durch ein entsprechendes Modell nachvollziehbar. Durch das Aufzeigen der beteiligten Ursache-Wirkungszusammenhänge wird einsichtig, was passiert. Das trifft sowohl für einen Einzelfall wie den Flugzeugabsturz zu als auch für einen allgemeinen Prozess. Ein Beispiel für Letzteres ist der Entstehungsmechanismus der typischen AIDS Symptome. Die untenstehende Abbildung zeigt ein vereinfachtes Modell auf.

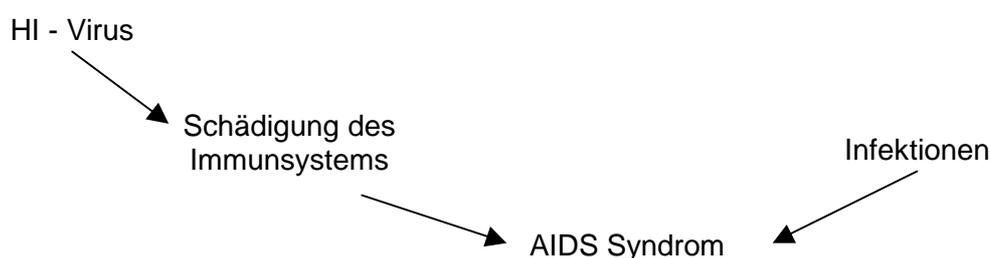


Abb. 1.3: Kausalmodell der Entstehung von AIDS

Nach diesem Modell führt das HI-Virus zu einer Schädigung des Immunsystems. Die Schädigung des Immunsystems führt mit anderen Infektionen zusammen zu dem bekannten Syndrom. Durch dieses einfache Modell konnten Phänomene erklärt werden, die für die Forscher bisher unverständlich geblieben waren. So war es lange Zeit für die Forscher mysteriös, welche verschiedenartigen Symptome bei AIDS Kranken auftraten. Diese waren für einen einzelnen Virus viel zu heterogen. Zudem wurden die äußeren Symptome durch andere Erreger verursacht, keines konnte auf den Virus selbst zurückgeführt werden. Selbst der Tod der Patienten war nicht durch den HI-Virus bedingt. Erst nachdem klar war, dass der Virus das Immunsystem schädigte, wurden diese Phänomene verständlich (Evans, 1993).

Eine zweite Funktion von Kausalmodellen ist es, Sachverhalte zu erklären. Sie zeigen den Weg auf, über den Ursachen zu ihren Effekten führen. Der Absturz der Concorde wird als Instantiierung eines Kausalmodells erklärt. Das Vorhandensein der Ursache (Reifenplatzen durch einen Gegenstand auf der Fahrbahn) und das Vorliegen aller Kausalzusammenhänge führte zu dem Absturz. Ebenso kann das Modell die vielen sicheren Starts von Concordes zuvor erklären. In diesen Fällen platzte kein Reifen oder der geplatze Reifen führte nicht zu Lecks in den Tragflächentanks. Das Kausalmodell von AIDS erklärt fast alle für Viruskrankheiten untypischen Phänomene. Die Vielfalt der Symptome ist durch verschiedene Erreger bedingt, die bei einem intakten Immunsystem vom menschlichen Organismus normalerweise in Schach gehalten werden. Auch ist klar, weshalb die

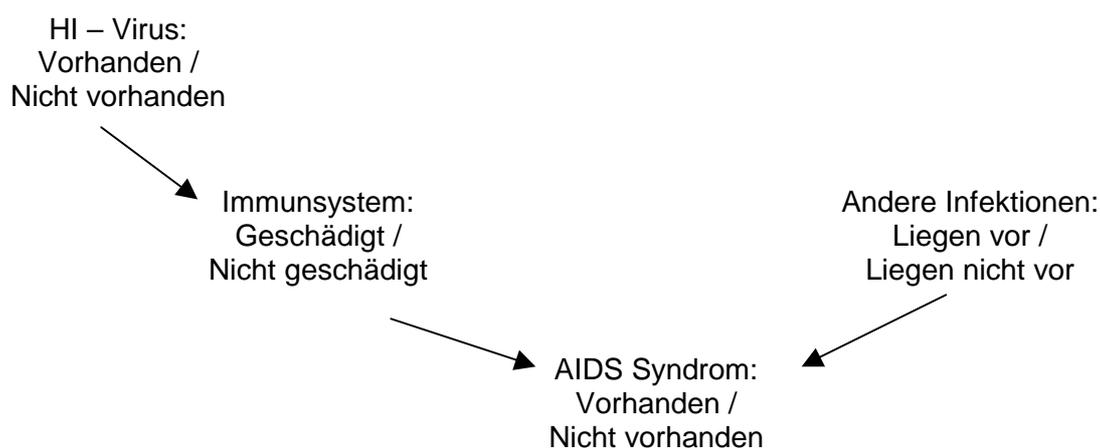
Erkrankten nicht an dem Virus selbst sterben. Andere Fragen bleiben aber weiterhin ungeklärt, z.B. weshalb es keine bestimmte Inkubationszeit gibt. Viren haben normalerweise einen festen Zeitraum, in dem sie nach der Infektion zu einer Erkrankung führen. Bei AIDS ist dieser extrem variabel, von Monaten bis zu Jahrzehnten. Dies deutet darauf hin, dass es noch weitere Faktoren geben muss, welche bisher unbekannt sind. Sobald diese entdeckt werden, wird das Kausalmodell entsprechend erweitert werden.

Ein überprüftes und an empirischen Daten bestätigtes Kausalmodells ermöglicht erfolgreiche Vorhersagen. So lässt das Absturzmodell die Vorhersage zu, dass bei einem erneuten Reifenplatzen eine erhebliche Gefährdung der Passagiere der Concorde bestehen würde. Für Passagiere anderer Flugzeuge ist die Gefahr deutlich geringer, da die Geschwindigkeiten beim Start niedriger sind und daher etwaige Reifenteile keine so große Wucht haben. Bei AIDS ist die bedauerliche, aber sichere Vorhersage, dass die Patienten nach Ausbruch der Krankheit an einer Vielzahl einfacher Infekte leiden werden.

Eine vierte Funktion von Kausalmodellen ist es, Handlungen zu leiten. Sie zeigen Möglichkeiten zur Intervention auf. Effekte können dadurch verhindert werden, dass das Auftreten der Ursache unterbunden wird oder dass weitere Ursachen eingeführt werden, die den Effekt hemmen. In Bezug auf AIDS zielen die Verhütungskampagnen darauf ab, die Infektion zu vermeiden. Medikamente dagegen wirken negativ auf die Erkrankung, d.h. sie wirken gegen die Symptome. Maßnahmen zur Vermeidung anderer Infektionen sind ebenfalls naheliegend und werden weithin praktiziert. Eine erste konkrete Maßnahme im Falle des Flugzeugabsturzes war es, die Intervalle zu verkürzen, in denen die Startbahnen auf herumliegende Teile abgesucht werden. Kausalmodelle leiten auch Handlungen, mit denen gewünschte Effekte herbeigeführt werden. Durch das gezielte Setzen von Ursachen können gewünschte Effekte hervorgerufen werden. So führt die Gabe abgeschwächter oder abgetöteter Viren dazu, dass der Körper Antikörper zur Abwehr gegen die Viren bildet. Dieser Vorgang ist sehr erwünscht, da er zu einer Immunisierung führt. Der Vorgang würde auch bei einer normalen Infektion auftreten, allerdings wäre er dann mit einer Erkrankung verbunden. Leider kommt es bei HI-Viren nicht zu einer Immunisierung.

Welche Bestandteile konstituieren ein Kausalmodell?

Ein Kausalmodell setzt sich aus Repräsentationen von Ereignissen und Kausalverbindungen zusammen. Die Ereignisse haben dabei den Status von Variablen, die verschiedene Ausprägungen annehmen können. Die umseitige Abbildung zeigt erneut das Kausalmodell für die Entstehung von AIDS, wobei die verschiedenen Ereignisausprägungen separat angeführt werden.



*Abb. 1.4: Kausalmodell der Entstehung von AIDS
Dichotome Ausprägungen der beteiligten Variablen*

Prinzipiell sind Kausalmodelle nicht auf dichotome Ereignisausprägungen beschränkt. So könnten anstelle des Vorhanden- bzw. Nichtvorhandenseins des AIDS Syndroms auch verschiedene Erkrankungsstufen (keine, leichte Infektionen, schwere Infektionen, Endstadium, Tod) als Ereignisausprägungen angeführt werden. Im weiteren Verlauf der Arbeit werden wir uns aber auf dichotome Ausprägungen beschränken.

Wie oben bereits erwähnt werden nur kausal relevante Ereignisse in ein Modell mit aufgenommen. Ein Ereignis ist dann kausal relevant, wenn sein Auftreten, die Wahrscheinlichkeit mindestens eines weiteren Ereignisses innerhalb des Modells verändert (Eells, 1991). Ausgeschlossen werden Ereignisse, die konstant an- oder abwesend sind, da sie die anderen Ereignisse nicht beeinflussen können.¹ Ebenfalls nicht mit aufgenommen werden unbekannte Faktoren.

Kausale Verbindungen sind gerichtete, mit einem bestimmten Vorzeichen versehene Zusammenhänge einer bestimmten Stärke. Sie werden durch die Pfeile repräsentiert. Die Richtung verläuft dabei stets von der Ursache hin zum Effekt. Damit kommt eine fundamentale Asymmetrie zum Ausdruck, Ursachen führen zu ihren Effekten, Effekte aber nicht zu ihren Ursachen (vgl. Waldmann, 1994, 1996). Das Vorzeichen bezieht sich darauf, dass die Ursache entweder den Effekt hervorruft (positive oder generative Ursache) oder den Effekt unterdrückt (negative oder inhibitorische Ursache). Kausalzusammenhänge

¹ Je nach Kontext fallen damit andere Ereignisse aus dem Modell heraus. So ist unter normalen Gegebenheiten auf der Erde überall Sauerstoff für eine Verbrennung vorhanden. Er kann daher bei der Modellbildung für die Verursachung von Feuer vernachlässigt werden. Im Weltraum ist nicht überall Sauerstoff vorhanden. In diesem Fall gehört er mit in ein Kausalmodell für einen Brand. Zur Erfassung dieses Unterschiedes wurde das Konzept des ‚focal set‘ (Cheng & Novick, 1992) entwickelt. Ereignisse, die innerhalb eines focal set konstant an- oder abwesend sind, können danach vernachlässigt werden.

innerhalb eines Modells sind keine deterministischen, sondern probabilistische Beziehungen. Die Pfeile bedeuten also nicht, dass das Ursacheereignis immer von einem Effekt ereignis gefolgt wird. Im Einzelfall kann es sein, dass nach einem Ursacheereignis kein Effekt sich einstellt, oder dass der Effekt ohne eine bekannte Ursache auftritt. Zum Beispiel kann bei AIDS auch Jahre nach einer Ansteckung mit dem HI-Virus keine Schädigung des Immunsystems vorliegen. Ebenso ist es möglich, dass das Immunsystem durch andere Erreger geschädigt wurde. Auch können viele typische AIDS Symptome wie Hepatitis oder Lungenentzündung auftreten, ohne dass eine HIV Infektion vorliegen muss.

Mit der Annahme, dass die Kausalbeziehungen innerhalb des Modells probabilistisch sind, ist aber nicht ausgeschlossen, dass die Kausalzusammenhänge in der Welt nicht doch deterministische Beziehungen sind (vgl. Pearl, 2000). Die unvollständige Wahrscheinlichkeitsbeziehung beruht nach dieser Konzeption darauf, dass nicht alle Ereignisse bekannt sind, die zu einem Effekt führen oder aber diesen verhindern können. Zum Beispiel sind gerade im medizinischen Bereich bei der Entstehung eines Krankheitsbildes stets sehr viele Faktoren relevant. Dies führt dazu, dass wahrscheinlich niemals ein vollständiges Kausalmodell gebildet werden kann. Daher macht es Sinn, sich bei der Modellbildung auf wesentliche Faktoren zu beschränken und dafür probabilistische Beziehungen in Kauf zu nehmen.

Die Stärke der Kausalzusammenhänge gibt an, wie wirksam die Ursache darin ist, den Effekt hervorzurufen. Hierfür gibt es eine ganze Reihe von Maßen. Diese werden im Abschnitt über quantitative Aspekte von Kausalmodellen im Einzelnen besprochen.

Welche Kausalmodelle gibt es?

Es gibt eine unendliche Vielzahl von Kausalmodellen. Diese lassen sich aber auf wenige einfache Grundmodelle zurückführen. Komplexe Kausalmodelle können durch die Verbindung dieser rekonstruiert werden. Die untenstehende Abbildung zeigt die drei Grundmodelle jeweils anhand eines Beispiels.

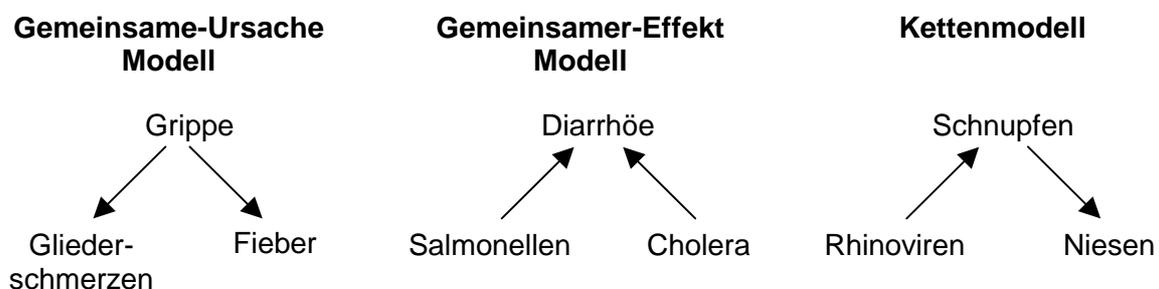


Abb. 1.5: Grundlegende kausale Strukturmodelle

Beim Gemeinsame-Ursache Modell führt eine einzelne Ursache zu mehreren Effekten. Dies ist ein typischer Fall bei Krankheiten. Eine Krankheit führt in der Regel zu mehreren verschiedenen Symptomen, die meistens alle zusammen auftreten. So verursacht eine Grippe nicht nur Fieber, sondern auch Gliederschmerzen. Bei einem Gemeinsamer-Effekt Modell führen mehrere verschiedene Ursachen zu einem gemeinsamen Effekt. Starke wässrige Durchfälle können sowohl eine Folge von Salmonellen als auch von Cholerafibrionen sein. Ein drittes grundlegendes Modell ist die kausale Kette. Rhinoviren führen zu dem allseits bekannten Schnupfen. Die Schnupfenerkrankung ist wiederum die Ursache für häufiges Niesen. Diese drei Modelle sind die einfachsten Strukturen, die aus mehreren Kausalzusammenhängen gebildet werden. Sie und ihre Eigenschaften werden uns im weiteren Verlauf der Arbeit beschäftigen.

Das komplexe Modell des Concorde-Absturzes setzt sich aus mehreren dieser einfachen Modelle zusammen. Den Anfang bildet ein Gemeinsamer-Effekt Modell bei dem die hohe Geschwindigkeit und das Platzen des Reifens zu einem Leck in den Tragflächentanks führen. Dieses Leck bildet zusammen mit den heißen Triebwerken ein weiteres Gemeinsamer-Effekt Modell mit dem Triebwerksbrand als gemeinsamen Effekt. Der Triebwerksbrand ist die Ursache innerhalb eines Gemeinsame-Ursache Modells mit dem Feuerschweif und dem Triebwerksausfall als Effekten. Der Triebwerksausfall ist wiederum Ausgangspunkt einer kausalen Kette, die beim Absturz endet.

Worin unterscheiden sich verschiedene Kausalmodelle?

Der Unterschied zwischen verschiedenen Modellen liegt in ihrer Struktur und den damit verbundenen Implikationen (vgl. Glymour & Cooper, 1999; Jensen, 1997; Pearl, 1988; Waldmann, 1996, 2000; Waldmann & Hagmayer, subm.). Nehmen wir das Gemeinsame-Ursache Modell. Was sollten wir beobachten, wenn dieses zutrifft? Wenn die Ursache gegeben ist, dann sollten die beiden mit ihr verbundenen Effekte auftreten. Beispielweise sollte ein Grippekranker sowohl Fieber als auch Gliederschmerzen haben. Wenn die Ursache nicht vorhanden ist, dann sollten auch die beiden Effekte ausbleiben. Ein Gesunder sollte also keines der beiden Krankheitssymptome aufweisen. Daraus folgt, dass ein statistischer Zusammenhang zwischen der Ursache und den beiden Effekten bestehen sollte. Was für einen Zusammenhang zwischen den beiden Effekten impliziert das Gemeinsame-Ursache Modell? Hängen die Effekte von einer gemeinsamen Ursache ab, dann ist zu erwarten, dass sie miteinander kovariieren. Fieber und Gliederschmerzen sollten als Symptome von Grippe häufig gemeinsam auftreten. Das Gemeinsame-Ursache Modell mit seiner Struktur impliziert also eine Abhängigkeit der beiden Effekte.

Liegt ein Gemeinsamer-Effekt Modell vor, dann sollten Zusammenhänge zwischen den beiden Ursachen und dem Effekt gegeben sein. Dabei sind aber verschiedene

Wirkungszusammenhänge möglich. So könnten beide Ursachen unabhängig voneinander auf den Effekt wirken. In diesem Fall sollte jede Ursache für sich einen statistischen Zusammenhang mit dem Effekt aufweisen. Dies trifft z.B. bei den beiden Krankheitserregern zu. Die negative Wirkung von Salmonellen ist nicht abhängig davon, ob gleichzeitig eine Infektion mit Cholerafibrionen vorliegt. Es könnte aber auch sein, dass die einzelnen Ursachen nur in Kombination miteinander wirken, sie also bei ihrer Wirkung interagieren. Dann sollte sich der Effekt nur zeigen, wenn beide Ursachen in einer bestimmten Ausprägung vorliegen. Ein statistischer Zusammenhang besteht in diesem Fall zwischen verschiedenen Ursachekonstellationen und dem Effekt. Dies traf z.B. beim Concorde-Absturz zu. Die Teile des geplatzten Reifens konnten nur wegen der sehr hohen Geschwindigkeit beim Start die Treibstofftanks beschädigen. Bei einer geringeren Geschwindigkeit wäre wahrscheinlich nichts passiert. Was für einen Zusammenhang zwischen den beiden Ursachen impliziert das Gemeinsamer-Effekt Modell? Dieses Modell impliziert keinen Zusammenhang zwischen den beiden Ursachen. Eine Infektion mit Cholera hat mit einer eventuellen Infektion mit Salmonellen nichts zu tun. Die beiden Ursachen sind unabhängig voneinander. Das heißt aber nicht, dass nicht beide doch zufällig kovariieren können. So ist es denkbar, dass das Wasser an einem bestimmten Ort mit beiden Erregern kontaminiert ist und so die Magen-Darm-Flora doppelt geschädigt wird. Allerdings ist diese Kovariation dann durch das verschmutzte Wasser als gemeinsame Ursache bedingt, sie rührt nicht daher, dass Salmonellen und Cholera beide zu Durchfall führen.

Das Kettenmodell führt wie die beiden anderen zu der Erwartung, dass zwischen den direkt kausal verbundenen Ereignissen ein Zusammenhang besteht. Wenn die erste Ursache gegeben ist, dann sollte das mittlere Ereignis auftreten. Wenn das mittlere Ereignis auftritt, dann sollte auch der Effekt am Ende der Kette auftreten. Rhinoviren und Schnupfen sowie Schnupfen und Niesen sollten also jeweils kovariieren. Wie sieht es mit dem Zusammenhang zwischen dem ersten Ereignis und dem letzten Effekt aus? Auch diese beiden sollten kovariieren. Wenn zwischen zwei Ereignissen ein über ein drittes Ereignis vermittelter Kausalzusammenhang besteht, dann sollte zwischen diesen beiden auch ein statistischer Zusammenhang bestehen. In unserem Beispiel ist also damit zu rechnen, dass zwischen Rhinoviren und Niesen ein Zusammenhang besteht. Dies ist auch tatsächlich der Fall. Für den Virus ist das Niesen seine besonders schnelle und effektive Art der Fortbewegung.

Die unten stehende Abbildung gibt nochmals einen Überblick über die Unterschiede und die Gemeinsamkeiten der drei Strukturen.

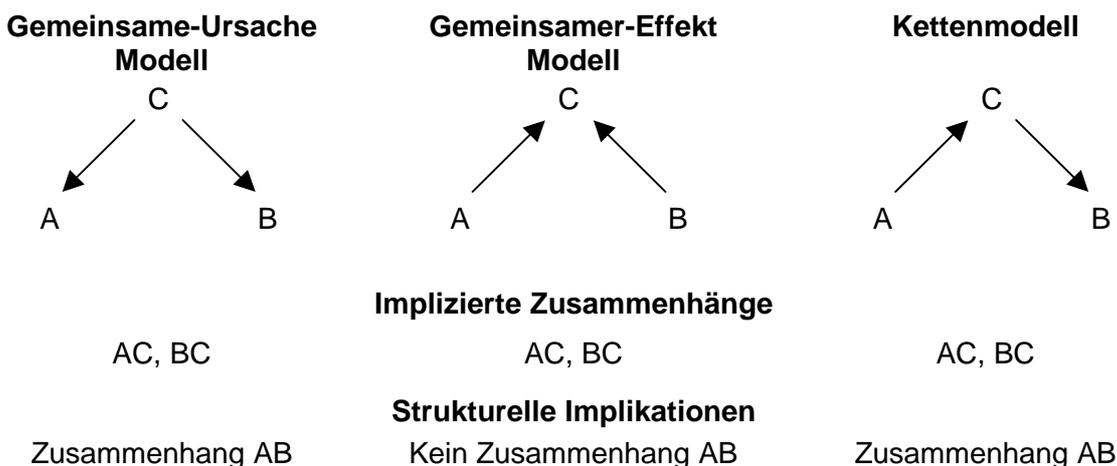


Abb. 1.6: Gemeinsamkeiten und Unterschiede zwischen den drei grundlegenden Kausalmodellen. Erläuterungen siehe Text.

Gemeinsam ist allen Strukturen, dass sie eine Abhängigkeit der direkt kausal verbundenen Ereignisse erwarten. Das heißt, alle Modelle sagen statistische Zusammenhänge zwischen A - C und B - C vorher. Der Unterschied liegt in den strukturellen Implikationen. Strukturelle Implikationen sind Vorhersagen über den Zusammenhang zwischen Ereignissen innerhalb eines Modells, die nicht über einen direkten Kausalpfeil verbunden sind. Bei den drei Grundmodellen sind dies die Ereignisse A und B. Das Gemeinsame-Ursache Modell und das Kettenmodell implizieren einen Zusammenhang dieser beiden Ereignisse. Wenn Ereignis A vorliegt, dann sollte auch Ereignis B vorliegen und umgekehrt. Dagegen impliziert das Gemeinsamer-Effekt Modell keine solche Kovariation. Dies trifft sich mit unserer Intuition, dass verschiedene Ursachen in der Regel nichts miteinander zu tun haben, dass aber verschiedene Effekte einer gemeinsamen Ursache immer alle auftreten sollten.

Wie entwickeln sich Kausalmodelle?

Kausalmodelle entwickeln sich häufig schrittweise. Zunächst wird Wissen über einzelne Kausalzusammenhänge erworben. Aufbauend auf diesem Wissen können die Modelle erweitert werden. Hierbei sind verschiedene Wege denkbar. Ein Ausgangspunkt können die beobachteten probabilistischen Zusammenhänge sein. Trotz des Vorliegens eines Kausalzusammenhangs werden immer wieder Fälle beobachtet, in denen der Effekt trotz Ursache ausbleibt oder der Effekt auftritt, ohne dass eine Ursache vorhanden war. Dies kann zu einer Suche nach anderen Ereignissen führen, die den Effekt ebenfalls bedingen. Werden solche entdeckt, wird das Modell um diese erweitert. So entsteht ein zunehmend komplexeres Gemeinsamer-Effekt Modell. Die vielfältigen Ursachefaktoren, die zu Krebserkrankungen führen, sind ein gutes Beispiel. Über die Jahre hinweg wurde erkannt,

dass u.a. genetische Prädispositionen, Viren, radioaktive Strahlung und karzinogene Substanzen das unkontrollierte Zellwachstum auslösen können. Ein zweiter Ansatzpunkt kann die Beobachtung weiterer, mit der Ursache kovariierender Ereignisse sein. So werden zusätzliche Nebeneffekte der Ursache erkannt und können in das Modell mit einbezogen werden. Es entwickelt sich ein Gemeinsame-Ursache Modell. Die von Contergan hervorgerufenen Geburtsschäden sind hierfür ein drastisches Beispiel. Eine dritte Möglichkeit besteht darin, den Prozess näher zu untersuchen, wie die Ursache zu ihrem Effekt führt. Dieser Ansatzpunkt ist typisch für wissenschaftliches Fragen. Ein solches Vorgehen resultiert in teilweise sehr komplexen kausalen Kettenmodellen. Ein Beispiel ist der in allen Biologielehrbüchern besprochene Zitronensäurezyklus. Diese Formen der Erweiterung können natürlich auch parallel zueinander ablaufen, so dass immer komplexere Kausalmodelle mit mehreren Ebenen und einer Vielzahl von Ursachen, vermittelnden Ereignissen und Effekten entstehen. Neben der Erweiterung von Kausalmodellen ist auch die Integration verschiedener unabhängiger Kausalrelationen ein Weg zu Bildung komplexer Kausalmodelle (vgl. Kapitel 2.4). So kann beispielsweise das Wissen, dass Aspirin zu Magenproblemen führt, mit dem später erworbenen Wissen, dass auch Bakterien zu Magenproblemen führen, zu einem Gemeinsamer-Effekt Modell kombiniert werden. Auch dieser Prozess führt zu immer größerer Komplexität.

Wissenschaftstheoretiker und Historiker haben untersucht, wie sich die Modelle in bestimmten Sparten der Wissenschaft entwickeln. Ein Beispiel für diese Entwicklung zeigen die beiden folgenden Kausalmodelle (vgl. Thagard, 1999). Sie beziehen sich auf die Entstehung von Zwölffingerdarmgeschwüren. Das einfache Kettenmodell auf der linken Seite zeigt den Wissensstand Mitte der achtziger Jahre. Zu dieser Zeit ging man in der Medizin noch davon aus, dass eine stressinduzierte Übersäuerung die Hauptursache von Magengeschwüren ist. Heutzutage wird von einem komplexen Modell ausgegangen (rechte Seite). Führende Mediziner sind der Ansicht, dass in der überwiegenden Mehrheit der Fälle (ca. 2/3) eine Infektion mit Bakterien die Ursache ist. Als eine weitere Ursache (ca. 20% der Fälle) hat sich eine Gruppe von Schmerzmitteln herausgestellt, zu der auch Aspirin zählt.

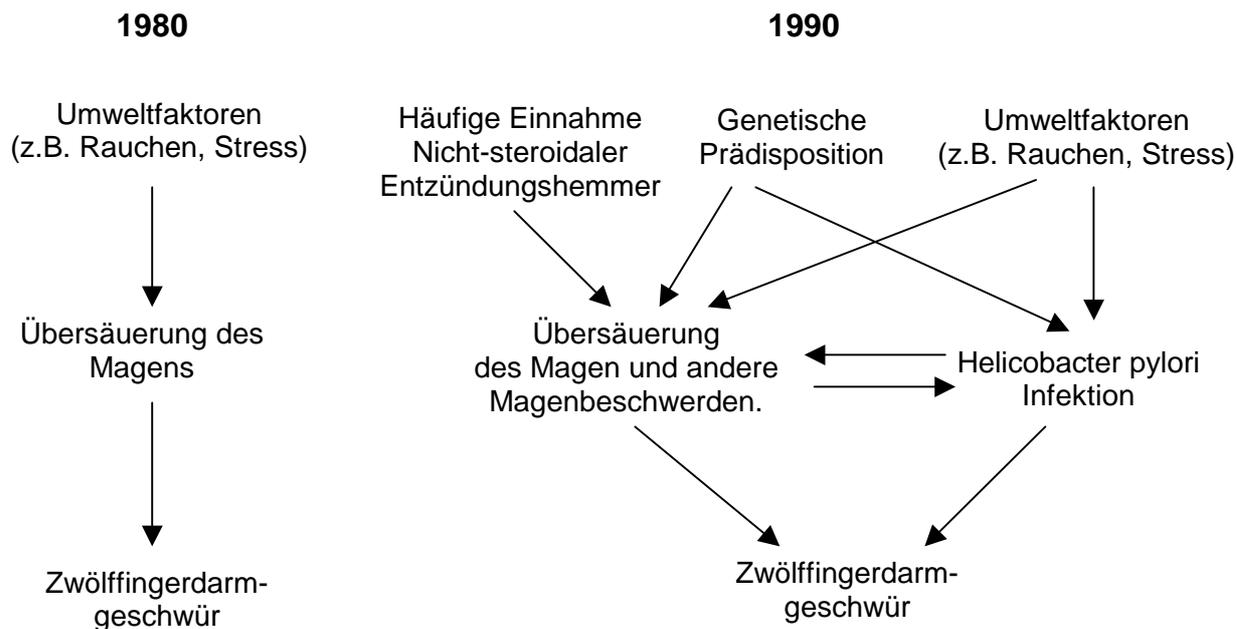


Abb. 1.7: Kausalmodelle der Entstehung von Zwölffingerdarmgeschwüren,
Linkes Modell: Stand Mitte 1980; Rechtes Modell: Mitte 1990
(verändert nach Thagard, 1999)

Fassen wir die wichtigsten qualitativen Aspekte von Kausalmodellen nochmals zusammen. Bei Kausalmodellen handelt es sich um die Repräsentationen von Ursache-Wirkungszusammenhängen. Sie setzen sich aus kausal relevanten Ereignissen und den Kausalzusammenhängen, welche diese Ereignisse verbinden, zusammen. Die Kausalzusammenhänge sind gerichtet und haben eine gewisse Stärke. Das wesentliche Merkmal von Kausalmodellen ist ihre Struktur und die mit dieser verbundenen Implikationen. Insbesondere anhand ihrer strukturellen Implikationen können verschiedene Kausalmodelle unterschieden werden. Die Struktur erklärt die beobachteten statistischen Zusammenhänge und lässt Vorhersagen sowohl für Einzelfälle wie für Gruppen von Fällen zu.

1.2 Formale Aspekte von Kausalmodellen²

Um Kausalmodelle genau definieren und effektiv einsetzen zu können ist es notwendig, diese formal zu fassen. Ziel dieses Teilkapitels ist es, zwei grundlegende Ansätze hierfür vorzustellen. Dabei werden wir uns einerseits mit linearen Strukturgleichungen und andererseits mit Bayesianischen Netzwerkmodellen beschäftigen. Diese Ansätze werden nur in Grundzügen vorgestellt, eine ausführliche Einführung ist im Rahmen dieser Arbeit nicht möglich.

1.2.1 Lineare Strukturgleichungen

Kausalmodelle lassen sich über lineare Strukturgleichungen formalisieren. Die Effektereignisse werden dabei als eine Funktion der Ursacheereignisse und der Stärke der beteiligten Kausalzusammenhänge definiert.

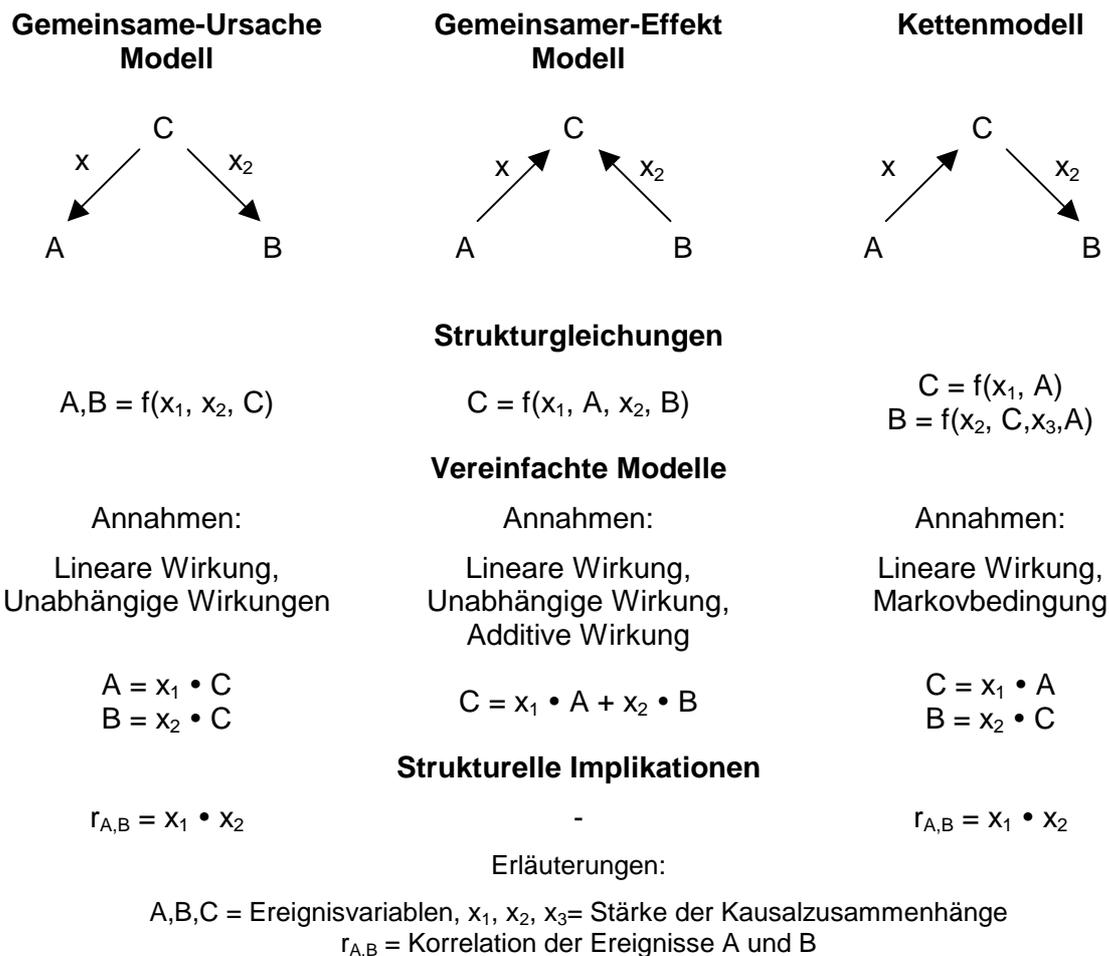


Abb. 1.8: Strukturgleichungen der grundlegenden Kausalmodelle

² Die Ausführungen in diesem Abschnitt beruhen auf folgenden Arbeiten: Glymour & Cooper, 1999; Jensen, 1997; Loehlin, 1998; Pearl, 1988, 2000; Spirtes et al., 1993.

Die allgemeinen Funktionen werden anwendbar, sobald zusätzliche Annahmen gemacht werden. In der Regel werden lineare Beziehungen angenommen. Die meisten heute gängigen Verfahren in der Statistik gehen von linearen Funktionen aus (Backhaus et al., 1994; Bortz, 1990). Nicht-lineare Beziehungen können häufig durch lineare erfolgreich angenähert werden. Um die Formeln noch weiter zu vereinfachen, müssen aber weitere Annahmen gemacht werden. Bei Gemeinsame-Ursache Modellen wird davon ausgegangen, dass die Wirkung der Ursache auf den einen Effekt unabhängig von der Wirkung auf den anderen Effekt ist. Diese Annahme ist plausibel. Weshalb sollte beispielsweise die schmerzlindernde Wirkung von Aspirin von der negativen Wirkung auf den Magen abhängen? Andere Mittel haben bei gleicher Wirksamkeit diesen Nebeneffekt nicht. Diese Annahme unabhängiger Wirkungen ist nicht gerechtfertigt, wenn von einer begrenzten Wirkmenge auszugehen ist. Ein Beispiel hierfür wäre die Aussendung einer bestimmten Menge an Transmittersubstanzen. Die Wirkung auf verschiedene Rezeptoren ist nicht unabhängig voneinander, da die zur Verfügung stehende Menge nur einmal verteilt werden kann. Für die meisten Sachverhalte macht aber die Annahme einer unabhängigen Wirkung Sinn. Unter diesen beiden Annahmen vereinfachen sich die Strukturgleichungen für die beiden Kausalzusammenhänge auf zwei einfache lineare Gleichungen, bei denen die Ursachevariable mit der Stärke der jeweiligen Kausalzusammenhänge gewichtet wird.

Beim Gemeinsamer-Effekt Modell werden üblicherweise drei Zusatzannahmen gemacht. Zunächst die Linearitätsannahme, die generell gemacht wird und daher unproblematisch ist. Zusätzlich vereinfacht die Annahme einer unabhängigen Wirkung der beiden Ursachen das Modell. Diese Annahme besagt, dass die Wirkung der einen nicht von der Wirkung der anderen Ursache abhängig ist bzw. dass keine Interaktion zwischen den beiden Ursachen vorliegt.³ Diese Annahme ist zum Beispiel dann verletzt, wenn der Effekt nur auftritt, wenn beide Ursachen gegeben sind. So führen beispielsweise die weit verbreiteten Herpesviren nur zu den bekannten Herpesbläschen, wenn das Immunsystem zusätzlich durch eine weitere Infektion belastet ist. Wird die Unabhängigkeitsannahme aber gemacht und ist zusätzlich davon auszugehen, dass die Wirkung der beiden Ursachen additiv ist, dann ergibt sich eine einfache Linearkombination der Wirkung.

Beim Kettenmodell kommt neben der Linearitätsannahme noch die sogenannte Markovbedingung hinzu. Diese besagt, dass jegliche Wirkung von A auf B über C vermittelt wird, $B = f(x_2, C) = f(x_2, C, x_3, A)$. Es liegt also keine Interaktion zwischen A und C in Bezug auf B vor. Dies ist auch das, was die graphische Darstellung eines kausalen Kettenmodells

³ Diese Annahme unabhängiger Wirkungen unterscheidet sich von der gleichnamigen Annahme für das Gemeinsame-Ursache Modell. Während beim Gemeinsame-Ursache Modell eine Abhängigkeit durch eine begrenzte Wirkmenge entstehen könnte, spielt diese bei Gemeinsamer-Effekt Modellen keine Rolle. Gemeinsam ist beiden Fällen, dass die Wirkung des einen Kausalzusammenhangs vom Vorliegen des anderen Kausalzusammenhangs beeinflusst wird.

vermittelt, A wirkt auf C und C wirkt auf B. A wirkt aber nicht direkt auf B. Ein Beispiel für die Verletzung dieser Annahme bieten genetische Prädispositionen. Diese erhöhen beispielsweise die Anfälligkeit für gewisse Erkrankungen, welche wiederum zu bestimmten Symptomen führen. Die Wirkung dieser Erkrankungen ist nun nicht als unabhängig von der genetischen Disposition zu sehen. Nur wenn diese vorliegt, führen die Krankheitserreger zu ihren Symptomen, ansonsten nicht. Inhaltlich ist die Markovbedingung überall dort gerechtfertigt, wo von klaren Ursache-Wirkungsketten ohne spezifische Interaktionen auszugehen ist. Der Vorteil dieser Zusatzannahme liegt in ihrer erheblichen Vereinfachung der Gleichung. Ansonsten wären bei jedem weiteren Kettenglied sämtliche möglichen Interaktionen aller vorherigen Ereignisse mit zu berücksichtigen.⁴

Diese oben genannten vereinfachenden Annahmen sind in der Regel notwendig für die Bestimmung von Strukturgleichungsmodellen. Ohne sie sind die Modelle häufig nicht identifizierbar. D.h. die Modellparameter lassen sich aufgrund der beobachteten Kovariationen nicht schätzen. Werden die allgemeinen Annahmen und die jeweiligen modellspezifischen Zusatzannahmen gemacht, dann können die Pfadgewichte als Maß für die Stärke des Kausalzusammenhangs interpretiert werden. Außerdem kann die Stärke der implizierten Zusammenhänge direkt über die Pfadkoeffizienten errechnet werden. So ist bei einem Gemeinsame-Ursache Modell die zu erwartende Korrelation der beiden Effekte gleich dem Produkt der beiden Pfadkoeffizienten. Die Höhe der Korrelation ist also direkt abhängig von der Stärke der Kausalzusammenhänge. Gleiches gilt für die Stärke der Korrelation zwischen der ersten Ursache und dem letzten Effekt bei einem Kettenmodell, auch hier gibt das Produkt der Pfadkoeffizienten die richtige Antwort. Über die Produkte der Pfadkoeffizienten der verbindenden Pfade kann auch bei sehr komplexen Modellen die zu erwartende statistische Abhängigkeit zweier weit entfernter Ereignisse ermittelt werden. Ob zwischen zwei weit entfernten Ereignissen eine statistische Abhängigkeit besteht und wie stark diese ist, dafür gibt es einfache, anhand des Graphen leicht zu überprüfende Kriterien (vgl. Wright's Rules, Loehlin, 1998, S.9).

⁴ Bei allen linearen Modellen müssen zusätzlich noch eine Reihe von Voraussetzungen erfüllt sein (vgl. Backhaus, 1994, S. 31 ff. für eine genaue Diskussion). So dürfen die Prädiktorereignisse nicht vollständig linear abhängig voneinander sein (Multikollinearität). Diese Voraussetzung besagt, dass die verschiedenen Ursachen eines Effektes möglichst unabhängig sein sollten. Genauso dürfen die Fehler, d.h. die unbekanntes weiteren Ursachen, nicht miteinander kovariieren (keine Autokorrelation). Eine dritte Voraussetzung ist, dass die bekannten nicht mit den unbekanntes Ursachen eines Effektes korrelieren sollten. Sind diese Voraussetzungen nicht erfüllt, dann werden die Schätzungen zunehmend unzuverlässig.

Vorteil der Repräsentation von Kausalmodellen über lineare Strukturgleichungsmodelle ist, dass sie gut mit kontinuierlichen Ursachen und Effekten umgehen können. Sie lassen zudem eine Überprüfung des angenommenen Kausalmodells zu. Ist das Modell korrekt, so müssen zwischen allen indirekt kausal verbundenen Ereignissen bestimmte statistische Abhängigkeiten bestehen. Die Stärke dieser Abhängigkeiten kann über die Höhe der Pfadkoeffizienten der beteiligten Kausalzusammenhänge genau vorhergesagt werden. Die Abweichung dieser aufgrund der Modellstruktur vorhergesagten Abhängigkeiten von den tatsächlich beobachteten liefert ein Maß für die Güte des Modells (vgl. Loehlin, 1998, S.54 ff.).

Der Nachteil von Strukturgleichungen ist deren inhärente Symmetrie. Jede Gleichung kann mathematisch gesehen beliebig in äquivalente umgeformt werden. Eine kausale Interpretation ist aber nur für eine bestimmte Form der Gleichung gegeben. Diesen Sachverhalt kann man sich leicht am zweiten Newtonschen Gesetz veranschaulichen. Nach diesem Grundgesetz der Mechanik ist die Kraft gleich dem Produkt von Masse und Beschleunigung ($f = m \cdot a$). Kausal gesehen macht diese Aussage keinen Sinn, die Kraft entsteht nicht aus beschleunigter Masse. Die Ursache ist in diesem Fall die an einer Masse angreifende Kraft, die Wirkung ist die daraus resultierende Beschleunigung. Trotzdem wird dieses physikalische Naturgesetz in obiger Form in den Lehrbüchern vermittelt. Das Problem der äquivalenten Gleichungen kann man sich auch über die Wirkung von Interventionen klar machen (vgl. Pearl, 2000, S. 346 ff.). Nehmen wir das kausale Kettenmodell mit $C = x_1 \cdot A$ und $B = x_2 \cdot C$. Nun setzen wir das vermittelnde Ereignis $C = 1$. Was folgt daraus? Vorausgesetzt wir wissen, dass die obigen Gleichungen ein Kausalmodell bilden, dann ändert sich durch diese Intervention lediglich der Wert von B auf $B = x_2$. Wissen wir nicht, dass die Gleichungen kausal interpretiert werden können, dann müssen wir auch den Wert von A neu bestimmen. Dass die Setzung des Ereignisses C keinen Einfluss auf die davor liegende Ursache A hat, kann also aus den Gleichungen selber nicht abgelesen werden. Diese Information vermittelt nur das Modell selbst. Nehmen wir zum Beispiel die kausale Kette die von Alkoholkonsum über vermindertes Reaktionsvermögen hin zu Verkehrsunfällen führt. Nun reduzieren wir das Reaktionsvermögen durch die Gabe sedierender Drogen. Damit steigt die Wahrscheinlichkeit für einen Verkehrsunfall deutlich an. Was sich aber nicht verändert ist der Alkoholkonsum. Dieser wird von der Einnahme eines Sedativums nicht beeinflusst.

Welche Schlussfolgerungen können wir aus dieser Diskussion ziehen? Strukturgleichungsmodelle sind eine gute Möglichkeit zur Repräsentation von Kausalmodellen. Um ihre Interpretierbarkeit sicherzustellen, ist es aber notwendig, die freie, äquivalente Umformung der Gleichungen zu beschränken.

1.2.2 Bayesianische Modelle⁵

Neben Strukturgleichungsmodellen sind Bayesianische Modelle eine zweite Alternative für die formale Repräsentation von Kausalmodellen. Bayesianische Modelle beruhen gleichermaßen auf der Wahrscheinlichkeits- wie auf der Graphentheorie. Die Wahrscheinlichkeitsrechnung liefert die Grundlage für Berechnungen, die Graphentheorie ermöglicht es, die Richtung der Kausalbeziehungen und die strukturellen Implikationen zu berücksichtigen. Ein Bayesianisches Netzwerkmodell wird wie folgt definiert:

- Ein Bayesianisches Netzwerk ist ein gerichteter azyklischer Graph. Azyklisch bedeutet, dass es keinen gerichteten Pfad gibt, der ein Ereignis mit sich selbst verbindet.
- Der Graph setzt sich zusammen aus einer Menge von Ereignissen und gerichteten Verbindungen zwischen diesen Ereignissen.
- Die Ereignisse haben eine begrenzte Menge von einander ausschließenden Ausprägungen.
- Jedem der Ereignisse mit seinen spezifischen Ursachen werden bedingte Wahrscheinlichkeiten $P(\text{Ereignis} \mid \text{Ursachen})$ zugeordnet.

(vgl. Jensen, 1998, S.18; Pearl, 2000, S. 13)

Die unten stehende Abbildung zeigt ein Beispiel für einen solchen Graphen. Die konstituierenden Ereignisse sind die Bakterien- und Vireninfektion, die Erkältung, die Schleimbildung und das Niesen. Die Verbindungen werden durch die Pfeile angegeben. Die Ereignisse können verschiedene Ausprägungen annehmen, z.B. ‚vorhanden‘ und ‚nicht vorhanden‘ oder ‚keine‘, ‚leicht‘ und ‚stark‘. Der Graph ist wie die Definition verlangt azyklisch. Zyklisch wäre er beispielsweise, wenn die Schleimbildung auf die Bakterieninfektion zurückwirken würde.

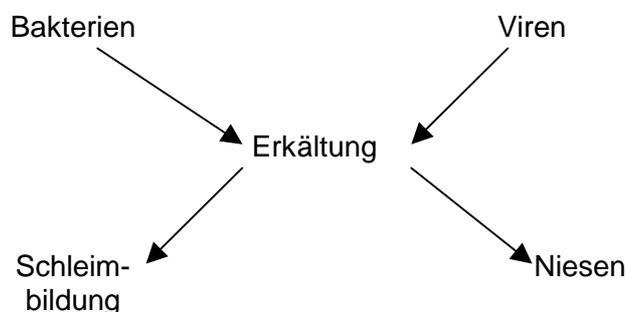


Abb. 1.9: Beispielgraph für ein Bayesianisches Netzwerkmodell

⁵ Dieser und der folgende Abschnitt setzen Grundwissen in Wahrscheinlichkeitsrechnung voraus. Für eine kurze Einführung in die wichtigsten Axiome und Theoreme siehe Jensen, 1997, Kapitel 2.

Kernstück des Modells sind die bedingten Wahrscheinlichkeiten, welche die Stärke der Kausalzusammenhänge erfassen. Diese gehören wie der Graph zu einem vollständigen Modell hinzu. Die Ursachen der Erkältung sind die Viren und die Bakterien, daher ist die bedingte Wahrscheinlichkeitsverteilung $P(\text{Erkältung} \mid \text{Viren, Bakterien})$ relevant. Die Ursache der Schleimbildung ist die Erkältung ($P(\text{Schleimbildung} \mid \text{Erkältung})$) und auch das Niesen wird von der Erkältung hervorgerufen ($P(\text{Niesen} \mid \text{Erkältung})$). Diese drei bedingten Wahrscheinlichkeitsverteilungen und der Graph bilden zusammen das Bayesianische Modell der dargestellten Kausalzusammenhänge. Wichtig dabei ist, dass es sich bei den bedingten Wahrscheinlichkeiten nicht um einzelne bedingte Wahrscheinlichkeiten für eine bestimmte Ausprägung der Ereignisse handelt, sondern jeweils um die Verteilung aller möglichen Kombinationen von Ereignisausprägungen. Für den Zusammenhang von Bakterien und Viren zur Erkältung heißt dies, dass

$$P(\text{Erkältung} \mid \text{Viren, Bakterien}) = \{P(\text{erkältung} \mid \text{viren.bakterien}), \\ P(\sim\text{erkältung} \mid \text{viren.bakterien}), P(\text{erkältung} \mid \sim\text{viren.bakterien}), \\ P(\sim\text{erkältung} \mid \sim\text{viren.bakterien}), P(\text{erkältung} \mid \text{viren.}\sim\text{bakterien}), \\ P(\sim\text{erkältung} \mid \text{viren.}\sim\text{bakterien}), P(\text{erkältung} \mid \sim\text{viren.}\sim\text{bakterien}), \\ P(\sim\text{erkältung} \mid \sim\text{viren.}\sim\text{bakterien})\}.$$

Um den Unterschied kenntlich zu machen, werden im folgenden für Ereignisvariablen Großbuchstaben und für spezifische Ausprägungen Kleinbuchstaben verwendet.

Wie sieht nun die Bayesianische Formulierung der drei grundlegenden Kausalmodelle aus? Die Abbildung auf der folgenden Seite gibt einen Überblick. Das jeweilige Bayesianische Modell gibt an, wie über die bedingten Wahrscheinlichkeiten zwischen Ursachen und Effekten und über die Basisraten der Ursachen die Wahrscheinlichkeitsverteilung über alle Ereignisvariablen bestimmt werden kann. Diese Ausgangsmodelle ergeben sich direkt aus den Axiomen der Wahrscheinlichkeitstheorie. Nach der sogenannten Kettenregel ist die gemeinsame Wahrscheinlichkeitsverteilung aller Ereignisvariablen gleich dem Produkt aller bedingter Wahrscheinlichkeiten innerhalb des Bayesianischen Netzwerkes. Diese Regel lässt sich auch auf Ereignisse und ihre Ursachen anwenden.

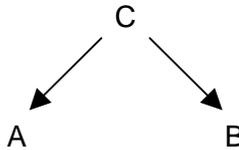
$$P(A_x) = \prod_i P(A_i \mid U_{A_i})^6$$

Dabei steht A_x für die Menge aller Ereignisvariablen und $P(A_x)$ für deren gemeinsame Wahrscheinlichkeitsverteilung. U_{A_i} ist die Menge der direkten Ursachen von A_i (vgl. Jensen, 1997, S. 20, Theorem 2.1; Pearl, 2000, S.14, Definition 1.2.1).

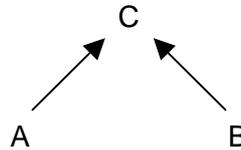
⁶ Für den Beweis dieses Theorems siehe Jensen, 1997, S. 20.

Dadurch ergeben sich folgende Formalisierungen für die drei grundlegenden Kausalmodelle.

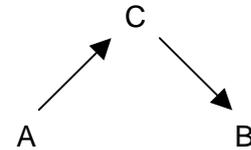
Gemeinsame-Ursache Modell



Gemeinsamer-Effekt Modell



Kettenmodell



Bayesianische Modelle

$$P(A,B,C) = P(A,B|C) \cdot P(C)$$

$$P(A,B,C) = P(C|A,B) \cdot P(A,B)$$

$$P(A,B,C) = P(B|A,C) \cdot P(C|A) \cdot P(A)$$

Vereinfachte Modelle

Annahmen:

Unabhängige Wirkungen

$$P(A,B,C) = P(A|C) \cdot P(B|C) \cdot P(C)$$

Annahmen:

Noisy-or Bedingung:
unabhängige Inhibitoren

$$P(A,B,C) = [1 - (1 - P(C|A)) \cdot (1 - P(C|B))] \cdot P(A) \cdot P(B)$$

Annahmen:

Markovbedingung:
Die Wirkung von C auf B ist
unabhängig von A

$$P(A,B,C) = P(B|C) \cdot P(C|A) \cdot P(A)$$

Erläuterungen:

A,B,C = Ereignisvariablen; $P(A,B,C)$ = Wahrscheinlichkeitsverteilung der Ereigniskombinationen;
 $P(A|C)$ = Verteilung der bedingten Wahrscheinlichkeiten von A und C,
 $P(C)$ = Basisraten des Ereignisses C

Abb. 1.10: Bayesianische Modelle für die grundlegenden kausalen Strukturmodelle

Gemeinsame-Ursache Modell

Beim Gemeinsame-Ursache Modell wird von einer unabhängigen Wirkung der einen Ursache auf ihre beiden Effekte ausgegangen. Nach der Definition der bedingten Wahrscheinlichkeit ist

$$P(A,B|C) = P(A|B,C) \cdot P(B|C)$$

Wirkt nun C unabhängig auf beide Effekte, dann gilt, dass

$$P(A|B,C) = P(A|C)$$

D.h., die bedingte Wahrscheinlichkeit von A auf C ist unabhängig von der Ausprägung von B. Auf obiges Beispiel angewandt bedeutet diese Gleichung, dass die Wahrscheinlichkeit für Niesen bei Vorliegen einer Erkältung und erhöhter Schleimbildung genau so hoch ist die Wahrscheinlichkeit für Niesen, wenn die Erkältung gegeben ist. Der Einfluss der Erkältung auf das Niesen ist also unabhängig von der Schleimbildung.

Unter dieser Bedingung resultiert das vereinfachte Modell

$$P(A,B|C) = P(A|C) \cdot P(B|C).$$

Die Wahrscheinlichkeit für Niesen und Schleimbildung bei einer Erkältung ist gleich dem Produkt der Wahrscheinlichkeit für Niesen, wenn eine Erkältung vorliegt und der Wahrscheinlichkeit für Schleimbildung bei einer Erkältung. Das bedeutet, dass innerhalb dieses Modells nur bedingte Wahrscheinlichkeiten, die für einzelne Kausalzusammenhänge stehen, verwendet werden.

Aus dem formalen Modell lassen sich auch die strukturellen Implikationen des Modells ableiten. Strukturelle Implikationen sind Vorhersagen über den Zusammenhang zwischen nicht direkt kausal verbundenen Ereignissen, in diesem Fall also der Zusammenhang der beiden Effekte. Das Gemeinsame-Ursache Modell impliziert eine Abhängigkeit der beiden Effekte untereinander, $P(B|A) \neq P(B)$. Eine Person, die niest, hat wahrscheinlich auch eine erhöhte Schleimsekretion. Das Modell hat aber auch noch eine zweite Implikation. Ist die Wirkung der Ursache auf den einen Effekt unabhängig von der Wirkung auf den zweiten Effekt, dann gilt wie oben gezeigt, dass $P(B|A,C) = P(B|C)$. Diese Gleichung besagt, dass konditional auf C die beiden Effekte A und B unabhängig voneinander sind. Konditional unabhängig bedeutet dabei, dass durch das Konstanthalten der gemeinsamen Ursache der Zusammenhang der beiden Effekte verschwindet. Betrachten wir also nur die an einer Erkältung Erkrankten, so haben von den Niesern genauso viele eine erhöhte Schleimabsonderung wie von den Personen, die nicht niesen. Kurz gesagt, das Gemeinsame-Ursache Modell impliziert eine unbedingte Abhängigkeit und eine bedingte Unabhängigkeit der Effekte.

Gemeinsamer-Effekt Modell

Rufen zwei Ursachen einen Effekt hervor, dann gilt, dass die Wahrscheinlichkeitsverteilung über alle Ereignisse gleich dem Produkt aus der bedingten Wahrscheinlichkeit des Effektes bei Vorliegen der Ursachen und der Wahrscheinlichkeit der Ursachen ist.

$$P(A,B,C) = P(C|A,B) \cdot P(A,B). \quad (1)$$

Dieser Fall entspricht im obigen Beispiel der Verursachung einer Erkältung durch Bakterien und Viren.

Eine Reihe von Annahmen führen zu einer erheblichen Vereinfachung des obigen Modells. Diese werden unter der Bezeichnung ‚Noisy or‘ zusammengefasst (Pearl, 1988, S.184 ff). Als erstes wird angenommen, dass die beiden Ursachen jeweils den Effekt immer hervorrufen. Die Beziehungen sind also deterministisch. Dass der Effekt dennoch in einigen Fällen ausbleibt, ist auf die Wirkung unbekannter Faktoren zurückzuführen, welche die Wirkung der Ursache verhindern. Diese inhibitorischen Faktoren sind spezifisch für jede der beiden Ursachen. In unserem Beispiel hieße das, dass sowohl Bakterien als auch die Viren immer zu einer Erkältung führen, es sei denn sie werden durch die jeweiligen Antikörper

ausgeschaltet. Die entscheidende Annahme ist nun, dass diese Inhibitoren unabhängig voneinander sind. Ob Antikörper für die Viren vorliegen, steht in keinem Zusammenhang dazu, ob Antikörper für die Bakterien vorhanden sind. Mit diesen Annahmen lässt sich nun das Modell wie folgt vereinfachen. Die Wahrscheinlichkeit, dass der Effekt durch den Inhibitor unterdrückt wird, obwohl die Ursache A vorliegt, ist

$$P(\sim C|A) = 1 - P(C|A).$$

Da die beiden Inhibitoren unabhängig voneinander sind, gilt

$$\begin{aligned} P(\sim C|A,B) &= P(\sim C|A) \cdot P(\sim C|B) \\ &= (1-P(C|A)) \cdot (1-P(C|B)) \text{ und} \\ P(C|A,B) &= 1-(1-P(C|A)) \cdot (1-P(C|B))^7 \end{aligned} \quad (2)$$

Durch Einsetzen von (2) in die Modellgleichung (1) ergibt sich

$$P(A,B,C) = [1-(1-P(C|A)) \cdot (1-P(C|B))] \cdot P(A,B). \quad (3)$$

Als nächstes gilt es die Wahrscheinlichkeitsverteilung der Ursachenkombinationen $P(A,B)$ zu vereinfachen. Da die beiden Ursachen innerhalb des Modells nicht kausal verbunden sind und auch keine gemeinsame Ursache haben, ist davon auszugehen, dass diese unabhängig sind, d.h. $P(A|B) = P(A)$. Daraus folgt, dass

$$P(A,B) = P(A|B) \cdot P(B) = P(A) \cdot P(B). \quad (4)$$

Aus (4) und (3) resultiert das vereinfachte Modell

$$P(A,B,C) = [1-(1-P(C|A)) \cdot (1-P(C|B))] \cdot P(A) \cdot P(B).$$

Dieses verwendet wie das Gemeinsame-Ursache Modell wiederum nur einfache bedingte Wahrscheinlichkeiten, die für einzelne Kausalzusammenhänge stehen, und Basisraten der ersten Ursachen.

Die vorgestellte ‚noisy-or‘ Annahme ist sehr schwach. Es wird lediglich verlangt, dass die Inhibitoren unabhängig sind. Das ist ein großer Vorteil dieser Annahme. Der Nachteil ist, dass sie zugegebenermaßen unanschaulich ist. Eine etwas stärkere Annahme führt einer anschaulicheren Ableitung des vereinfachten Modell. Gehen wir dazu erneut von unserem Beispiel aus. Wenn bei einer Person sowohl eine Virus- wie eine Bakterieninfektion vorliegt, dann kann die Erkältung entweder durch die Viren, die Bakterien oder durch beide ausgelöst worden sein. Formal gesprochen, die Wahrscheinlichkeit, dass eine Erkältung bei Viren oder Bakterien auftritt, ist gleich der Summe der Wahrscheinlichkeit, dass eine Erkältung durch Viren hervorgerufen wird und der Wahrscheinlichkeit, dass eine Erkältung durch Bakterien

⁷ Diese Gleichung gilt nur, wenn keine weiteren Ursachen den Effekt hervorrufen, d.h. $P(c|\sim a.\sim b) = 0$. Ist dies nicht der Fall, so muss das Modell entsprechend erweitert werden.

verursacht wird, minus der Wahrscheinlichkeit, dass es durch beide gleichzeitig hervorgerufen wird.⁸

$$P(\text{Erkältung} | \text{Viren oder Bakterien}) = P(E|V) + P(E|B) - P(E|V.B)$$

Nehmen wir nun an, dass Bakterien und Viren unabhängig voneinander die Erkältung auslösen, dann ergibt sich die Wahrscheinlichkeit der Erkältung, wenn eine doppelte Infektion vorliegt als das Produkt der einzelnen bedingten Wahrscheinlichkeiten.

$$P(E|V.B) = P(E|V) \cdot P(E|B)$$

Durch Einsetzen in obigen Gleichung ergibt sich

$$P(E|V,B) = P(E|V) + P(E|B) - P(E|V) \cdot P(E|B)$$

Diese Gleichung ist äquivalent zu (2)

$$\begin{aligned} P(C|A,B) &= 1 - [1 - P(C|A)] \cdot (1 - P(C|B)) \\ &= 1 - [1 - P(C|A) - P(C|B) + P(C|A) \cdot P(C|B)] \\ &= P(C|A) + P(C|B) - P(C|A) \cdot P(C|B) \end{aligned}$$

Diese zweite Ableitung des vereinfachten Gemeinsamer-Effekt Modells zeigt zweierlei. Als erstes belegt sie, dass dieses Modell unter verschiedenen starken Vorannahmen resultiert. Zweitens liefert sie eine klare Interpretation des vereinfachten Modells. Das Gemeinsamer-Effekt Modell besagt, dass die Wirkung jeder Ursache nicht von der Wirkung der anderen Ursachen innerhalb des Modells abhängt. Viren rufen unabhängig von Bakterien eine Erkältung hervor und umgekehrt.

Folgende strukturelle Implikationen ergeben sich aus dem Modell für den Zusammenhang der beiden Ursachen. Wie oben gezeigt geht das Gemeinsamer-Effekt Modell in seiner vereinfachten Form von einer Unabhängigkeit der Ursachen aus, $P(B|A) = P(B)$. Diese Annahme ist durch das Fehlen einer weiteren gemeinsamen Ursache der Ursachen gerechtfertigt. Was passiert aber, wenn auf C konditionalisiert wird? Dann gilt, dass $P(B|A,C) \neq P(B|C)$, sprich A und B sind konditional auf C abhängig voneinander. Wenn der Effekt gegeben ist und bereits eine der Ursachen vorliegt, dann ist es wahrscheinlicher, dass die andere Ursache nicht zusätzlich gegeben ist. Diese Implikation des Modells wird auch als das ‚principle of explaining away‘ bezeichnet (Jensen, 1997, S. 12). Hat eine Patientin bereits eine Virusinfektion und dadurch bedingt eine Erkältung, dann hat sie wahrscheinlich nicht auch noch zusätzlich einen bakteriellen Infekt. Dies trifft zu, obwohl in der Gesamtbevölkerung bakterielle und virale Infektionen nicht miteinander korreliert sind.

⁸ Die Wahrscheinlichkeit, dass die Erkältung bei beiden Infektionen ausgelöst wird, muss subtrahiert werden, da diese bereits Bestandteil jeder der beiden anderen Wahrscheinlichkeiten ist und sonst doppelt gezählt würde. Diesen Umstand kann man sich leicht an einem Venn-Diagramm klarmachen.

Kettenmodell

Das ursprüngliche Kettenmodell lautet

$$P(A,B,C) = P(B|A,C) \cdot P(C|A) \cdot P(A).$$

Es gilt also $P(B|A,C)$ zu vereinfachen. Dazu wird beim Kettenmodell angenommen, dass die Wirkung von C auf B unabhängig von A ist. Zum Beispiel ist anzunehmen, dass eine Erkältung zu Niesen führt, unabhängig davon, ob sie nun durch Viren oder Bakterien hervorgerufen wurde. Ist diese sogenannte Markovbedingung erfüllt, dann ist A bedingt auf C unabhängig von B:

$$P(B|A,C) = P(B|C)$$

Daraus resultiert das vereinfachte Modell.

$$P(A,B,C) = P(B|C) \cdot P(C|A) \cdot P(A)$$

Bei diesem werden erneut nur einfache bedingte Wahrscheinlichkeiten und die Basisrate der ersten Ursache verwendet. Inhaltlich besagt dieses Kettenmodell, dass der gesamte Einfluss von A auf B über C vermittelt wird, oder mit anderen Worten, A kann B nicht beeinflussen, wenn C konstant gehalten wird. Rhinoviren führen nicht zu Niesen, wenn der Schnupfen durch Medikamente kuriert wird.

Die strukturellen Implikationen des Kettenmodells betreffen den Zusammenhang zwischen der ersten Ursache und dem letzten Effekt. Für diese beiden Ereignisse impliziert das kausale Kettenmodell eine Abhängigkeit, $P(B|A) \neq P(B)$. Eine Infektion mit Rhinoviren sollte mit häufigem Niesen einher gehen. Konditional auf das intermittierende Ereignis C sind die Ereignisse aber unabhängig voneinander, $P(B|A,C) = P(B|C)$. Wird das vermittelnde Ereignis konstant gehalten, dann kann sich die erste Ursache nicht auf den letzten Effekt in der Kette auswirken. Wird eine sich anbahnende Erkältung medikamentös unterdrückt, dann niest die Person trotz eines Infektes nicht häufiger als andere Personen. Die strukturellen Implikationen des Kettenmodells stimmen damit mit denen des Gemeinsame-Ursache Modells überein. Beide implizieren eine unbedingte Abhängigkeit von A und B, aber eine bedingte Unabhängigkeit. Das Gemeinsamer-Effekt Modell impliziert jedoch genau das Gegenteil, eine unbedingte Unabhängigkeit und eine bedingte Abhängigkeit.

Komplexe Kausalmodelle

Über Bayesianische Modelle lassen sich beliebig komplexe Modelle abbilden. So ergibt sich für die Wahrscheinlichkeitsverteilung unseres Erkältungsbeispiels folgendes Modell (die Ereignisse wurden durch ihre Anfangsbuchstaben abgekürzt):

$$P(V,B,E,S,N) = P(N|E) \cdot P(S|E) \cdot P(E|V,B) \cdot P(V) \cdot P(B)$$

Dieses komplexe Modell ergibt sich also direkt aus den drei einfachen Grundmodellen.

Über die gesamte Wahrscheinlichkeitsverteilung können alle Formen von Kausalbeziehungen abgebildet werden. D.h., es können auch nicht lineare Zusammenhänge und Interaktionen erfasst werden. Allerdings müssen bei einer Repräsentation über diese Wahrscheinlichkeitsverteilung bei n Ereignissen 2^n Wahrscheinlichkeiten bestimmt und gespeichert werden. Jede denkbare Kombination von Ereignisausprägungen ist dann einzigartig und muss unabhängig von allen anderen betrachtet werden. Das Kausalmodell des Concorde-Absturzes bezieht beispielsweise 10 Ereignisse mit ein, d.h., es müssten 1024 Wahrscheinlichkeiten gespeichert werden. Dieses Vorgehen ist ineffektiv und überschreitet selbst bei moderaten Ereigniszahlen schnell jegliche Kapazitätsgrenzen. Werden aber die entsprechenden vereinfachenden Annahmen gemacht, dann brauchen nur die bedingten Wahrscheinlichkeiten die Kausalzusammenhänge betreffend und die Basisraten der ersten Ursachen gespeichert werden. Beim Concorde-Beispiel sind dies 9 Kausalrelationen bei drei Initialursachen. Das bedeutet, dass das Kausalmodell nur 18 bedingte Wahrscheinlichkeiten und 3 Basisraten umfasst. Über das Modell mit seinen strukturellen Implikationen können mit diesen wenigen Angaben alle 1024 Wahrscheinlichkeiten der gemeinsamen Wahrscheinlichkeitsverteilung bestimmt werden.

Vergleicht man die Bayesianischen Modelle mit den linearen Strukturgleichungsmodellen, so zeigen sich ähnliche Vor- und Nachteile. Die nicht vereinfachten Bayesianischen Modelle haben keine spezifische kausale Interpretation. Allein durch die Vereinfachung gewinnen sie ihre Interpretierbarkeit. Dabei wird der Graph, welcher integraler Bestandteil des Kausalmodells ist, für die Modellspezifikation genutzt. Die vereinfachten Modelle berücksichtigen die Kausalrichtung. Ursachen erscheinen bei den bedingten Wahrscheinlichkeiten als Bedingungen, Effekte als bedingte Ereignisse. Der Nachteil der Bayesianischen Kausalmodelle ist, dass eine Vereinfachung nur unter den oben vorgestellten Zusatzannahmen legitimiert und erfolversprechend ist.

Die Formalisierung von Kausalmodellen über lineare Strukturgleichungsmodelle und über Bayesianische Netzwerkmodelle sind zwei alternative Wege. Diese widersprechen sich aber nicht. Es konnte vielmehr gezeigt werden, dass sie in einander überführbar sind (Pearl, 2000). Unter den entsprechenden Annahmen stimmen auch die strukturellen Implikationen beider vollständig überein. Wenn wir daher im Folgenden nur noch Bayesianische Kausalmodelle betrachten, so gelten die gemachten Aussagen auch für lineare Strukturgleichungsmodelle.

1.3 Quantitative Aspekte von Kausalmodellen

Wenn wir nach der Stärke eines Kausalzusammenhangs fragen, so wollen wir wissen, wie häufig durch die untersuchte Ursache der in Frage stehende Effekt hervorgerufen wird. Eine Beobachtung des Zusammenhangs der beiden Ereignisse zeigt uns aber nur, wie häufig Ursache und Effekt zusammen auftreten. Wir müssen also aus den beobachteten statistischen Zusammenhängen erst auf Kausalität schließen. In der Literatur wurden verschiedene statistische Indikatoren für die Wirksamkeit einer Ursache vorgeschlagen. Der erste Abschnitt dieses Teilkapitels stellt drei dieser Maße für Kausalität vor. Im zweiten Abschnitt wenden wir uns wieder den grundlegenden Kausalmodellen zu. Für die Bestimmung der Stärke eines Kausalzusammenhangs ist es nämlich entscheidend, in welches Kausalmodell er eingebettet ist. Je nach umgebender Kausalstruktur sind andere statistische Zusammenhänge indikativ für Kausalität. In Abschnitt drei wird dann beispielhaft vorgeführt, wie die Wahrscheinlichkeitsverteilung der grundlegenden Kausalmodelle aus den bedingten Wahrscheinlichkeiten der Kausalzusammenhänge berechnet werden kann. Der vierte Abschnitt führt diese Berechnung weiter und zeigt, wie die strukturellen Implikationen der Modelle bestimmt werden können.

1.3.1 Bestimmung der Stärke einzelner Kausalzusammenhänge

Kontingenz

Ein erster Ansatz beruht auf einem probabilistischen Konzept von Kausalität. Nach dieser Auffassung ist eine Ursache ein Ereignis, welches die Wahrscheinlichkeit des Auftretens eines anderen Ereignisses (ihres Effektes) verändert (Eells, 1991; Salmon, 1980). Ein Medikament reduziert die Auftretenswahrscheinlichkeit einer Erkrankung und eine Infektion erhöht diese. Sowohl das Medikament als auch die Infektion würden wir daher als kausal wirksam anerkennen. Aufbauend auf diesem Grundverständnis wird die Stärke eines Kausalzusammenhangs über die Kontingenz (ΔP) bestimmt. Diese ist definiert als die Differenz aus der Wahrscheinlichkeit des Effektes E, wenn die Ursache U gegeben ist ($P(e|u)$), und der Wahrscheinlichkeit des Effektes, wenn die Ursache nicht gegeben ist ($P(e|\sim u)$).

$$\Delta P = P(e|u) - P(e|\sim u) \quad (\text{Definition der Kontingenz})$$

Ruft eine Ursache einen Effekt hervor, dann sollte der Effekt häufiger nach der Ursache als ohne die Ursache auftreten. Die Kontingenz für eine generative Ursache sollte daher positiv

sein. Verhindert eine Ursache einen Effekt, dann sollte dieser seltener nach der Ursache als ohne diese Ursache eintreten. Die Kontingenz ist in diesem Fall negativ.

Die Kontingenz ist aber nur unter bestimmten Bedingungen ein valides Maß für Kausalität. So besteht beispielsweise tatsächlich eine positive Kontingenz zwischen dem Auftreten von Störchen und der Geburtenrate. Trotzdem mag niemand auf einen Kausalzusammenhang schließen. Der Grund für die positive Kontingenz ist, dass sowohl das Auftreten von Störchen als auch die Geburtenrate vom Wetter abhängig sind. Das Wetter ist die gemeinsame Ursache innerhalb eines Gemeinsame-Ursache Modells. Die positive Kontingenz der beiden Ereignisse ist nichts anderes als der von dieser Struktur implizierte Zusammenhang. Die Kontingenz zwischen zwei Ereignissen ist daher nur ein valider Indikator für Kausalität, wenn keine solche gemeinsame Ursache vorliegt. Die entsprechende Annahme wird als Exogenität bezeichnet und wird uns noch mehrfach begegnen⁹.

Die Kontingenz ist das in der Psychologie am häufigsten verwendete Maß für die kausale Wirksamkeit (Cheng & Novick, 1990, 1992; Jenkins & Ward, 1965; Shanks et al., 1996; Waldmann, 1994, 1996; Ward & Jenkins, 1965; Wasserman et al. 1993). Auch eine deterministische Auffassung von Kausalität ist mit der Kontingenz als Maß vereinbar. Probabilistische Zusammenhänge sind gemäß dieser Auffassung die Folge von unbekanntem Ursachen. Ist die Wahrscheinlichkeit des Effektes, wenn die Ursache nicht vorliegt größer als Null ($P(e|\sim u) > 0$), dann kommt das daher, dass noch andere unbekannte Ursachen den Effekt hervorrufen. Entsprechend kann es auch sein, dass die Ursache nicht immer von ihrem Effekt gefolgt wird ($P(e|u) < 1$). Dies ist darauf zurückzuführen, dass in den Ausnahmefällen noch weitere, den Effekt hemmende Ursachen gegeben sind.

Der große Vorteil der Kontingenz ist es, dass diese direkt aus den Beobachtungsdaten bestimmt werden kann. Die beobachtete Kontingenz zwischen einem Ursacheereignis und seinem nachfolgenden Effekt ereignis gibt nach diesem Ansatz direkt die Wirkung der Ursache wieder. Der Nachteil der Kontingenz ist es, dass sie nicht zwischen Kausalzusammenhängen und beobachtbaren statistischen Zusammenhängen differenziert. Da sich die Kontingenz mit jeder weiteren Beobachtung ändert, würde sich auch der Kausalzusammenhang ständig ändern. Außerdem gibt es Fälle, in denen die Kontingenz nicht den Kausaleinfluss wiederzugeben scheint. Untersuchen wir die Wirkung einer Ursache auf einen Effekt und stellen fest, dass der Effekt immer nach der Ursache auftritt ($P(e|u) = 1,0$), er aber auch ohne die Ursache immer gegeben ist ($P(e|\sim u) = 1,0$), dann ist die Kontingenz $\Delta P = 0$. D.h., wir sollten zu dem Schluss gelangen, dass die Ursache den Effekt

⁹ Wenn zu erwarten ist, dass keine Exogenität vorliegt, dann kann diese durch zufällige Manipulation der Ursache hergestellt werden (Pearl, 2000). Angenommen C beeinflusst A und B. Wir wollen aber wissen, ob A auch B beeinflusst. Durch das experimentelle Setzen von A wird dieses Ereignis unabhängig von C, $P(A|C) \rightarrow P(A)$. C beeinflusst dann nur noch die Wahrscheinlichkeit von B. Damit liegt aber Exogenität vor und die Kontingenz wird wieder interpretierbar.

nicht beeinflusst. Dieser Schluss scheint aber kontraintuitiv zu sein (vgl. Wu & Cheng, 1999). Wir haben eher den Eindruck, dass wir in diesem Fall keine Aussage über die Wirksamkeit der in Frage stehen Ursache machen können. Wegen dieser Nachteile wurden andere Maße für Kausalität vorgeschlagen.

Kausale Power

Ein anderes Maß für kausale Wirksamkeit ergibt sich, wenn wir von einem Konzept „kausaler Power“ (Cartwright, 1989; Cheng, 1997) ausgehen. Nach dieser Auffassung steht hinter jedem beobachteten Kausalzusammenhang eine Ursache, die eine Wirkung (Power) auf den Effekt ausübt (Buehner & Cheng, 1997; Cheng, 1997; Wu & Cheng, 1999). Die Power ist dabei ein theoretisches Konstrukt zur Erklärung der beobachteten Zusammenhänge. Sie muss aus den Beobachtungen erst erschlossen werden. Eine ähnliche Idee liegt auch mechanistischen Ansätzen zur Kausalität zugrunde (Ahn et al., 1995; Ahn & Bailenson, 1996). Für diese sind beobachtbare Zusammenhänge die Folge dahinter verborgen liegender Mechanismen. Die kausale Power ist definiert als diejenige Wahrscheinlichkeit, mit der eine Ursache durch ihr Auftreten einen Effekt hervorruft, wenn keine weiteren Ursachen gegeben sind. Da in realen Kontexten aber immer weitere unbekannte Ursachen gegeben sind, ist die kausale Power nur eine theoretische, praktisch nicht direkt bestimmbare Größe. Sie lässt sich aber aufgrund der beobachtbaren Zusammenhänge unter gewissen Zusatzannahmen schätzen. So geht in die beobachtbaren bedingten Wahrscheinlichkeiten neben der Wirkung der in Frage stehenden Ursache U stets auch die Wirkung anderer Ursachen A mit ein. Unter der Annahme, dass diese alternativen Ursachen alle den Effekt fördern, können diese als eine generative Ursache betrachtet werden. Eine zweite Annahme, welche für die Bestimmung kausaler Power gemacht werden muss, ist, dass die beiden Ursachen U und A unabhängig voneinander auf den Effekt wirken.¹⁰ Ist dies der Fall, dann gilt für die bedingten Wahrscheinlichkeiten (vgl. Cheng, 1997, S.373):

$$P(e|u) = p_u + P(a|u) \cdot p_a - p_u \cdot P(a|u) \cdot p_a \quad (1)$$

$$P(e|\sim u) = P(a|\sim u) \cdot p_a \quad (2)$$

Dabei steht p_u für die Power der Zielursache und p_a für die Power der alternativen Ursache. Die Wahrscheinlichkeit des Effektes nach Eintreten der Ursache ist die Summe aus der Power der Ursache und der Wirkung der alternativen Ursache, sofern sie ebenfalls aufgetreten ist, vermindert um die Schnittmenge aus beiden Wirkungen. Die Wahrscheinlichkeit des Effektes, wenn die Ursache nicht eingetreten ist, ergibt sich als Produkt aus der bedingten Wahrscheinlichkeit, dass bei Fehlen der Ursache trotzdem die alternative Ursache aufgetreten ist und deren Power. Gehen wir nun, wie bei den

¹⁰ Der Ansatz lässt sich auch auf Interaktionen erweitern (Cheng & Novick, in Vorbereitung)

Gemeinsamer-Effekt Modellen üblich, von der Unabhängigkeit der untersuchten Ursache von der Alternativursache aus, dann ist $P(a|u) = P(a|\sim u) = P(a)$. Damit vereinfachen sich die beiden Gleichungen auf

$$P(e|u) = p_u + P(a) \cdot p_a - p_u \cdot P(a) \cdot p_a \quad (1')$$

$$P(e|\sim u) = P(a) \cdot p_a \quad (2')$$

Die Kontingenz ΔP ergibt sich als Differenz von (1') und (2') als

$$\begin{aligned} \Delta P &= p_u + P(a) \cdot p_a - p_u \cdot P(a) \cdot p_a - P(a) \cdot p_a \\ &= p_u - p_u \cdot P(a) \cdot p_a \end{aligned}$$

Löst man diese Gleichung nach der Power der Zielursache auf, dann erhält man:

$$p_u = \Delta P / (1 - P(a) \cdot p_a)$$

Nach Gleichung (2') ist $P(a) \cdot p_a = P(e|\sim u)$. Damit ergibt sich die kausale Power einer Ursache als

$$p_u = \Delta P / (1 - P(e|\sim u)) \quad (\text{Bestimmung der kausalen Power})$$

Kausale Power als Maß für die Wirksamkeit einer Ursache beruht also ebenfalls auf der Kontingenz, berücksichtigt aber zusätzlich den Einfluss weiterer Ursachen und korrigiert das Maß um diesen Anteil.¹¹ Die Intuition, die hinter dieser Korrektur steht, ist folgende: Wenn der Effekt bereits häufig aufgrund anderer Ursachen eintritt, kann die Zielursache ihre Wirkung nur selten zeigen. Die Wirkung der Zielursache wird nur in den Fällen sichtbar, in denen der Effekt nicht schon durch andere Ursachen hervorgerufen wurde. Tritt der Effekt zum Beispiel in 8 von 10 Fällen ohne die Zielursache auf, dann hat die Zielursache bei 10 Beobachtungen nur in 2 Fällen die Chance ihre Wirkung zu zeigen. Wird der Effekt also in 9 von 10 Fällen nach Auftreten der Ursache beobachtet, so heißt dies, dass die Zielursache in einem der beiden möglichen Fälle, in denen sie ihre Wirkung zeigen konnte, diese auch gezeigt hat. Ihre Power ist also 0,5. Nichts anderes ergibt sich nach der Bestimmungsgleichung für die Power: $p_u = (0,9 - 0,8) / (1 - 0,8) = 0,1 / 0,2 = 0,5$. Die Kontingenz ist in diesem Fall sehr schwach mit $\Delta P = 0,1$. Der Ursache sollte nach diesem Maß also nur ein geringer Einfluss auf den Effekt zugeschrieben werden.

Vorteil der kausalen Power ist es, dass sie zwischen den beobachteten Zusammenhängen und den dahinter stehenden Kausalzusammenhängen unterscheidet. Durch die sich aus der Wahrscheinlichkeitsrechnung zwingend ergebende Bestimmungsgleichung für die Power kann auf den alleinigen Einfluss der Zielursache zurückgeschlossen werden. Ein weiterer Vorteil ist, dass die Bestimmungsgleichung für kausale Power natürliche Randbedingungen besitzt. Wir hatten im vorherigen Abschnitt den Fall besprochen, in dem der Effekt sowohl nach der Ursache als auch ohne diese stets auftritt, ($P(e|u) = P(e|\sim u) = 1$). Die Kontingenz beträgt in diesem Fall Null. Die Ursache wirkt

¹¹ Für inhibitorische Zielursachen kann eine entsprechende Ableitung gemacht werden. Diese resultiert in folgender Gleichung (Cheng, 1997, S.375): $p_u = -\Delta P / P(e|\sim u)$.

nach diesem Maß also nicht auf den Effekt. Die Bestimmungsgleichung für Power ist in diesem Fall unbestimmt, $p_u = 0 / 0$. D.h., der Ansatz kausaler Power sagt in diesem Fall, dass eine Beurteilung der Wirksamkeit der Ursache nicht möglich ist. Dieses Ergebnis entspricht meines Erachtens eher unserer Intuition. Für das Kontingenzmaß liegen keine solchen Randbedingungen vor. Nachteil der Power ist, dass sie im Vergleich zu Kontingenzen erheblich schwerer zu bestimmen ist. Während eine Kontingenz durch Beobachten und Lernen der statistischen Zusammenhänge direkt gegeben ist, muss für die Power eine zusätzliche Schlussfolgerung gezogen werden. Kontingenz und Power setzen beide Exogenität voraus. Ist der beobachtete Zusammenhang auf eine nicht bekannte gemeinsame Ursache zurückzuführen, dann liefern beide Maße eine Fehleinschätzung.

Probability of Sufficiency and Necessity

Bevor wir handeln, überlegen wir, ob unsere Handlung den gewünschten Effekt haben wird. Wir wollen also vor dem Hintergrund, dass weder die Ursache noch der Effekt sich bisher ereignet haben, vorhersagen, mit welcher Wahrscheinlichkeit der Effekt auftreten wird, wenn die Ursache gesetzt würde. Haben wir beispielsweise starke Kopfschmerzen, dann überlegen wir, mit welcher Wahrscheinlichkeit die Einnahme eines bestimmten Schmerzmittels uns Besserung bringen wird. Dies ist die Frage danach, wie hinreichend die Ursache für den Effekt ist. Die ‚Probability of sufficiency‘ (Pearl, 1999) gibt diese Wahrscheinlichkeit an. Sie wird definiert als die Wahrscheinlichkeit, dass die Ursache U den Effekt E hervorrufen wird, gegeben, dass weder die Ursache noch der Effekt in dem konkreten Fall bereits aufgetreten ist. Formal bedeutet dies

$$PS = P(E, U | \sim e, \sim u) \quad (\text{Probability of sufficiency})$$

Dieses Maß gibt direkt an, wie stark über eine Handlung ein Effekt beeinflusst werden kann. Es konnte gezeigt werden, dass diese Wahrscheinlichkeit durch Beobachtung bestimmt werden kann, wenn zwei Zusatzannahmen gemacht werden (Pearl, 1999). Die erste ist die bereits bekannte Annahme von Exogenität, welche besagt, dass die Ursache und der Effekt keine gemeinsamen Ursachen haben. Die zweite Annahme ist die Monotonizität der Wirkung und besagt, dass das Auftreten der Ursache im Einzelfall den Effekt nur hervorrufen, niemals aber unterdrücken kann. Die Ursache ist also entweder generativ oder sie wirkt nicht. Bei Exogenität und Monotonizität der Ursache gilt:

$$PS = [P(e|u) - P(e|\sim u)] / [1 - P(e|\sim u)]$$

Damit entspricht aber die Bestimmung der ‚Probability of sufficiency‘ der Bestimmung für generative Power. Diese kann also nicht nur als Maß für die generelle kausale Wirksamkeit der Ursache, sondern auch als Maß für den Einfluss kausalen Handelns im Einzelfall genommen werden.

Neben der Frage, wie hinreichend eine Ursache für ihren Effekt ist, interessiert auch, wie notwendig die in Frage stehende Ursache für den Effekt ist. Diese Frage ist z.B. bei Schadensersatzklagen relevant. Folgendes Szenario wäre denkbar: Eine Familie verklagt den Betreiber eines benachbarten Atomkraftwerks auf Schmerzensgeld, weil ein Kind an Leukämie erkrankt ist. Die Ursache (Atomkraftwerk) und der Effekt (Leukämie) sind beide also schon eingetreten. Die Frage, die das Gericht an einen Gutachter stellt, ist, mit welcher Wahrscheinlichkeit das Kind nicht an Leukämie erkrankt wäre, wenn es kein Atomkraftwerk in der Nähe gegeben hätte. Ist diese Wahrscheinlichkeit hoch, d.h., ohne das Kraftwerk wäre das Kind heute noch gesund, dann hat die Familie Anspruch auf Schadenersatz. Die Wahrscheinlichkeit, dass der Effekt nicht aufgetreten wäre, wenn die Ursache ausgeblieben wäre, unter der Voraussetzung, dass beide sich bereits ereignet haben, gibt also an, wie notwendig die Ursache für den Effekt ist. Das zugehörige Maß wird als ‚Probability of necessity‘ bezeichnet und wird definiert als:

$$PN = P(\sim E, \sim U | e, u) \quad (\text{Probability of necessity})$$

Unter den beiden oben genannten Annahmen der Exogenität und der Monotonizität, kann diese Wahrscheinlichkeit aufgrund von Daten bestimmt werden als

$$PN = [P(e|u) - P(e|\sim u)] / P(e|u)^{12}$$

Beide Wahrscheinlichkeiten lassen sich auch in ein Maß vereinen. Dieses gibt an, wie notwendig und hinreichend die Ursache für den Effekt ist. Die ‚Probability of necessity and sufficiency‘ kann empirisch ermittelt werden über

$$PNS = P(e|u) - P(e|\sim u) \quad (\text{Probability of necessity and sufficiency})$$

Diese Formel ist identisch mit der Bestimmung der Kontingenz. Dennoch wäre es falsch zu sagen, dass die Kontingenz angibt, wie notwendig und hinreichend die Ursache für den Effekt ist. Diese besondere Interpretation ist nur unter den Zusatzannahmen zulässig. Die Annahme der Exogenität ist dabei sowohl für die Kontingenz als auch für die PNS wichtig. Liegt eine gemeinsame Ursache vor, so verliert auch die Kontingenz ihre Validität. Aber die Kontingenz setzt keine Monotonizität voraus. Es kann durchaus sein, dass zum Beispiel ein Medikament einem bestimmten Patienten schadet, obwohl es allen anderen hilft. Die Kontingenz als Maß für die kausale Wirksamkeit benötigt also schwächere Zusatzannahmen. Der Vorteil der beiden neuen Maße ist, dass sie Antworten auf für uns sehr interessante Fragen geben. Sie beantworten die Frage, wie stark ich durch ein Setzen der Ursache den Effekt beeinflussen kann (Probability of sufficiency), und wie relevant die Ursache für den beobachteten Effekt war (Probability of necessity).

¹² Diese Wahrscheinlichkeit entspricht rein formal bis auf das positive Vorzeichen, dem Maß für inhibitorische kausale Power (vgl. Fußnote 11 oder Cheng, 1997). Die Interpretationen beider Maße weichen aber deutlich voneinander ab. Die ‚Probability of necessity‘ gibt nicht an, wie stark die Ursache den Effekt verhindert, sondern, wie notwendig sie für ihn ist.

1.3.2 Bestimmung von Kausalzusammenhängen innerhalb eines Kausalmodells

Die oben vorgestellten Maße gelten zunächst nur für einfache Kausalzusammenhänge. Das bedeutet, es wird angenommen, dass nichts über andere Ereignisse, die in Zusammenhang mit der Zielursache oder dem Zieleffekt stehen, bekannt ist. Im folgenden Absatz wollen wir uns nun mit der Frage nach Maßen für kausale Wirksamkeit bei verschiedenen Kausalmodellen beschäftigen. Bei Kausalmodellen liegen stets weitere Ereignisse vor. Die Hauptfrage für die Bestimmung eines Kausalzusammenhangs ist dann, ob diese Zusatzereignisse für den beobachteten statistischen Zusammenhang zwischen den beiden Zielereignissen verantwortlich sein könnten. Ist dies der Fall, dann muss der Einfluss dieser Ereignisse kontrolliert werden. Mit anderen Worten, immer wenn eine mögliche Konfundierung droht, muss diese über geeignete Maßnahmen bei der Bestimmung des Kausalzusammenhangs ausgeschlossen werden. Kausalitätstheoretiker haben sich mit dieser Frage eingehend beschäftigt und sind zu folgender Empfehlung gelangt (vgl. Eells, 1991, Kapitel 3 und 4):

“When assessing a causal relation between an event C and an event E, we must hold fixed all nonsubsequent causal factors that are causally relevant to E (i.e. alternative causes or interacting events), and we must hold not fixed subsequent causal factors that are directly or indirectly caused by C.” (Eells, 1991, S. 206)

Diese Regel gibt für alle möglichen kausalen Strukturmodelle an, welches der normativ richtige Indikator für das Vorliegen und die Stärke eines Kausalzusammenhangs ist. Welche Anweisungen ergeben sich aus dieser normativen Regel für die drei grundlegenden Kausalmodelle? Die Abbildung 1.11 zeigt nochmals die drei Modelle und die Formeln zur Bestimmung der Kontingenzen und Powerwerte zwischen den kausal verbundenen Ereignissen. Da die ‚Probability of sufficiency and necessity‘ formal der Kontingenz entspricht, wird diese nicht gesondert aufgeführt.

Für Stärke des Kausalzusammenhangs indicative statistische Zusammenhänge

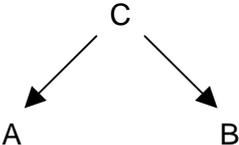
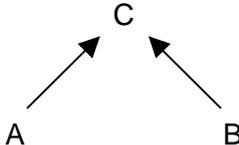
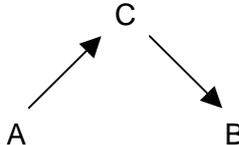
Gemeinsame-Ursache Modell	Gemeinsamer-Effekt Modell	Kettenmodell
		
Unbedingte Kontingenzen	Bedingte Kontingenzen bei generativen Ursachen	Unter Markovbedingung: Unbedingte Kontingenzen
$\Delta P_{CA} = P(a c) - P(a \sim c)$	$\Delta P_{AC} = P(c a.\sim b) - P(c \sim a.\sim b)$	$\Delta P_{AC} = P(c a) - P(c \sim a)$
$\Delta P_{CB} = P(b c) - P(b \sim c)$	$\Delta P_{BC} = P(c b.\sim a) - P(c \sim b.\sim a)$	$\Delta P_{CB} = P(b c) - P(b \sim c)$
Unbedingte Power bei generativen Ursachen	Bedingte Power bei generativen Ursachen	Unter Markovbedingung: Unbedingte Power bei generativen Ursachen
$p_{CA} = \Delta P_{CA} / 1 - P(a \sim c)$	$p_{AC} = \Delta P_{AC} / 1 - P(c \sim a.\sim b)$	$p_{AC} = \Delta P_{AC} / 1 - P(c \sim a)$
$p_{CB} = \Delta P_{CB} / 1 - P(b \sim c)$	$p_{BC} = \Delta P_{CA} / 1 - P(c \sim b.\sim a)$	$p_{CB} = \Delta P_{CB} / 1 - P(b \sim c)$

Abb. 1.11: Bestimmung der Stärke der Kausalzusammenhänge innerhalb eines Kausalmodells

Bei einem Gemeinsame-Ursache Modell werden die beiden Effekte A und B von der einen Ursache direkt hervorgerufen. Damit greift der zweite Teil der obigen Regel. Bei der Bestimmung des Einflusses von C auf A darf B nicht konstant gehalten werden, da B ein von C direkt verursachter Faktor ist. Entsprechendes gilt umgekehrt für den Zusammenhang von C und B, A darf hier bei der Bestimmung nicht konstant gehalten werden. Stress führt beispielsweise zu Magenproblemen und erhöhtem Blutdruck. Wollen wir nun bestimmen, wie stark Stress den Magen belastet, dann dürfen wir nicht den Blutdruck als weiteres Symptom konstant halten. Die Anweisung, die anderen Ereignisse nicht konstant zu halten, ist gleichbedeutend mit der Verwendung unbedingter Kontingenzen bzw. unbedingter Power als Maß für Kausalität. Unbedingt heißt also, dass bei der Bestimmung der Kontingenz die An- oder Abwesenheit des Zusatzereignisses nicht gesondert berücksichtigt wird.

Bei einem Gemeinsamer-Effekt Modell sind zwei Ereignisse gegeben, welche beide den einen Effekt beeinflussen. Wollen wir die Stärke des Kausalzusammenhangs zwischen A und C bestimmen, dann müssen wir nach dem ersten Teil der Regel B konstant halten. B ist ein nicht auf die Zielursache A folgender, aber für den Effekt C relevanter Kausalfaktor. Sollten wir diesen anderen Faktor besser konstant anwesend oder abwesend halten? Rufen sowohl die Zielursache als auch die zweite Ursache den Effekt hervor, dann sollte die

Zweitursache abwesend gehalten werden (Cheng & Novick, 1990, 1992). Verhindert die Zielursache eher den Effekt, dann muss mindestens eine andere generative Ursache konstant anwesend gehalten werden, damit die inhibitorische Wirkung überhaupt erkannt werden kann (vgl. Cheng, 1997). Ist nichts Genaueres über die Wirkungsrichtung der beiden Ursachen bekannt, so macht es Sinn, die zweite Ursache in einer Reihe von Fällen konstant anwesend und in einer Reihe anderer Fälle konstant abwesend zu halten. Ein Beispiel: Nikotin und Koffein führen beide zu einer Erregung des Sympathikus. Wollen wir nur die Wirkung des Nikotins bestimmen, dann müssen wir die Koffeineinnahme der Probanden kontrollieren. Am besten wäre es, wenn die Probanden während der Experimente kein Koffein zu sich nehmen würden. Das Konstanthalten der zweiten Ursache ist gleichbedeutend mit der Verwendung bedingter Kontingenzen bzw. bedingter Power als Maß für den Kausalzusammenhang. Die bedingte Kontingenz ist definiert als die Differenz der Wahrscheinlichkeit des Effektes bei Vorliegen der Ursache und der Wahrscheinlichkeit des Effektes bei Fehlen der Ursache unter der Bedingung, dass alle anderen Ursachen konstant gehalten werden. Bei zwei generativen Ursachen ergibt sich die bedingte Kontingenz als

$$\Delta P_{AC} = P(c|a.\sim b) - P(c|\sim a.\sim b).$$

Bei Verwendung von Power als Maß für Kausalität gilt

$$p_{AC} = \Delta P_{AC} / 1 - P(c|\sim a.\sim b).$$

Auch hier sollten nur die Fälle verwendet werden, in denen die jeweils andere Ursache abwesend ist.

Bei einem kausalen Kettenmodell ist die Anweisung für den ersten der beiden Kausalzusammenhänge klar. Das letzte Effektereignis darf als ein der ersten Ursache nachfolgendes, von dieser indirekt kausal abhängiges Ereignis nicht konstant gehalten werden. Für den zweiten Kausalzusammenhang hängt die Anweisung davon ab, ob es sich um eine Markovkette handelt oder nicht. Ist die Markovbedingung erfüllt, dann ist das erste Ursacheereignis konditional unabhängig vom letzten Effekt. Die erste Ursache ist damit ein nicht nachfolgendes und aufgrund der konditionalen Unabhängigkeit ein nicht kausal relevantes Ereignis für den Zieleffekt. In diesem Fall braucht es nicht konstant gehalten werden (für solche Fälle macht die obige Regel keine zwingenden Vorschriften). Ist die Markovbedingung aber nicht erfüllt, dann ist die erste Ursache ein nicht nachfolgendes, für den Zieleffekt indirekt kausal relevantes Ereignis und muss damit konstant gehalten werden. Zwei Beispiele hierzu: Übermäßiges Essen führt zur Fettleibigkeit, welche wiederum die Gelenke schädigt. Wollen wir den Einfluss von hoher Kalorienzufuhr auf Fettleibigkeit ermitteln, dann dürfen wir nicht die Gelenkschäden der Probanden mit einbeziehen. Wollen wir die Auswirkung der Fettleibigkeit auf Gelenkschäden untersuchen, dann ist es irrelevant, woher die Fettleibigkeit kommt. Ein Beispiel für einen Fall in dem die Markovbedingung nicht

erfüllt ist, sind Erbkrankheiten wie Mukoviszidose. Wollen wir den Auswirkung starker Schleimabsonderung auf die Atemwege untersuchen, dann ist es relevant, ob diese durch einen genetischen Defekt bedingt ist oder auf eine kurzfristige Infektion zurückzuführen ist. Bei einer Markovkette sind für beide Teilzusammenhänge die unbedingten Kontingenzen bzw. Powerwerte indikativ für die kausale Wirksamkeit. Ist keine Markovkette gegeben, dann ist für den ersten Zusammenhang die unbedingte Kontingenz und für den zweiten Zusammenhang die bedingte Kontingenz heranzuziehen. Entsprechendes gilt für die Power.

Weshalb ist es bei Kausalmodellen so wichtig, ob die bedingten oder unbedingten Kontingenzen das normativ richtige Maß für die Stärke des Kausalzusammenhangs ist? Die Antwort hierauf liefert das sogenannte Simpson Paradox (Cartwright, 1989; Eells, 1991; Simpson, 1951). Dieses besagt, dass ein statistischer Zusammenhang zwischen zwei Variablen in einer Population in allen Teilgruppen verschwinden oder sich vom Vorzeichen her sogar umdrehen kann. Mit anderen Worten, es ist möglich, dass die unbedingten Kontingenzen einen positiven Kausalzusammenhang anzeigen und die bedingten Kontingenzen einen negativen. Dieses Paradoxon ergab sich in einem berühmten Fall an der University of California in Berkeley (vgl. Cartwright, 1989). Untersucht wurde, ob Frauen bei der Zulassung zu den Graduiertenstudiengängen benachteiligt werden. Auf Universitätsebene wurden tatsächlich sehr viel mehr Männer als Frauen aufgenommen. Betrachtete man aber die einzelnen Departments, so zeigte sich, dass tendenziell sogar mehr Frauen als Männer angenommen wurden. Die untenstehende Abbildung veranschaulicht diesen Fall an einem konkreten Zahlenbeispiel.

	Department A		Department B		Universität	
	Annahme	Ablehnung	Annahme	Ablehnung	Annahme	Ablehnung
Männer	80	10	0	10	80	20
Frauen	10	0	10	80	20	80
Kontingenz	$\Delta P = -0,11$		$\Delta P = -0,11$		$\Delta P = +0,60$	

Abb. 1.12: Zahlenbeispiel für ein Simpson Paradox, die Zahlen geben Absolutwerte an; Die rechten beiden Spalten ergeben sich als Summe der linken vier Spalten

Insgesamt wurden von 100 Männern 80 aufgenommen, aber nur 20 von 100 Frauen. In Department A wurden alle 10 Frauen aufgenommen, die sich beworben hatten und 80 von 90 Männern, in Department B wurde keiner der 10 Bewerber akzeptiert, aber immerhin 10 der 90 Frauen. Werden Frauen nun diskriminiert oder bevorzugt? Diese Frage lässt sich nur mit Hilfe eines Kausalmodells beantworten. Wenn das Department ein kausal relevanter Faktor für die Annahme oder Ablehnung ist, dann muss nach Departments getrennt

ausgewertet werden und die bedingten Kontingenzen geben den Zusammenhang richtig wieder. Ist das Department nicht relevant, weil beispielweise die Bewerber per Zufall den Departments zugewiesen werden, dann dürfen sie nicht berücksichtigt werden und die Gesamtuniversität ist die richtige Referenzpopulation. Da die Departments selber die Auswahl durchführen und auch sehr unterschiedliche Annahmehquoten haben, ist das Department klar relevant. Demnach wurden die Frauen eher bevorzugt. Das Kausalmodell und die zugehörige normative Regel geben also an, welcher der statistischen Zusammenhänge den Kausalzusammenhang wiedergeben.

1.3.3 Berechnung von Kausalmodellen

Die Formalisierung von Kausalmodellen als Bayesianische Modelle ermöglicht die Berechnung der Wahrscheinlichkeitsverteilung aller Ereigniskombinationen über die Stärke der beteiligten Kausalzusammenhänge. Der folgende Abschnitt führt diese Berechnung am Beispiel der drei grundlegenden Kausalmodelle vor. Für alle drei Modelle wurden dabei dieselben Annahmen zu Grunde gelegt. Die Basisrate der Ursachen war stets $P(u) = 0,5$, die Wahrscheinlichkeit des Effektes bei Vorliegen der Zielursache in Abwesenheit aller anderen bekannten Ursachen $P(e|u) = P(e|u.\sim a) = 0,8$ und die Wahrscheinlichkeit des Effektes bei Abwesenheit aller anderen bekannten Ursachen $P(e|\sim u) = P(e|\sim u.\sim a) = 0,2$. Die bedingten Wahrscheinlichkeiten waren für beide Kausalzusammenhänge gleich.

Bei allen Modellen resultiert unter diesen Annahmen eine Kontingenz von $\Delta P = 0,6$ und eine Power von $p = 0,75$ für beide Kausalzusammenhänge. Die unbedingte Kontingenz und Power beim Gemeinsame-Ursache Modell und beim Kettenmodell stimmen dabei mit der bedingten Kontingenz und bedingten Power beim Gemeinsamer-Effekt Modell überein. D.h., die normativ richtigen Indikatoren für Kausalität sind in allen Modellen gleich. Für die Berechnung der Wahrscheinlichkeitsverteilungen spielt der Unterschied zwischen den Kontingenzen und Powerwerten keine Rolle, wichtig sind nur die bedingten Wahrscheinlichkeiten.

Der Vorgang der Berechnung der Wahrscheinlichkeitsverteilung soll anhand einer Ereigniskombination für jedes Modell exemplarisch vorgeführt werden. Beim Gemeinsame-Ursache Modell ergibt sich die Wahrscheinlichkeit, dass alle drei Ereignisse zusammen auftreten als

$$P(a.b.c) = P(a|c) \cdot P(b|c) \cdot P(c) = 0,8 \cdot 0,8 \cdot 0,5 = 0,32.$$

Beim Gemeinsamer-Effekt Modell gilt

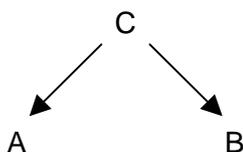
$$P(a.b.c) = [1-(1-P(c|a.\sim b)) \cdot (1-P(c|b.\sim a)) \cdot (1-P(c|\sim a.\sim b))^{13}] \cdot P(a) \cdot P(b) \\ = [1- (1-0,8) \cdot (1-0,8) \cdot (1-0,2)] \cdot 0,25 = 0,242.$$

Beim Kettenmodell gilt

$$P(a.b.c) = P(c|a) \cdot P(b|c) \cdot P(a) = 0,8 \cdot 0,8 \cdot 0,5 = 0,32.$$

Die Wahrscheinlichkeiten der anderen Ereigniskombinationen können auf gleichem Wege ermittelt werden.

Gemeinsame-Ursache Modell

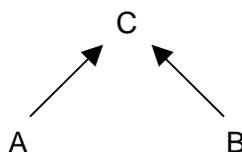


$$P(A,B,C) = P(A|C) \cdot P(B|C) \cdot P(C)$$

$$P(c) = P(\sim c) = 0,5 \\ P(a|c) = P(b|c) = 0,8 \\ P(a|\sim c) = P(b|\sim c) = 0,2$$

$$P(a.b.c) = 0,32 \\ P(a.b.\sim c) = 0,02 \\ P(a.\sim b.c) = 0,08 \\ P(\sim a.b.c) = 0,08 \\ P(a.\sim b.\sim c) = 0,08 \\ P(\sim a.b.\sim c) = 0,08 \\ P(\sim a.\sim b.c) = 0,02 \\ P(\sim a.\sim b.\sim c) = 0,32 \\ \Sigma P(A,B,C) = 1,0$$

Gemeinsamer-Effekt Modell



$$P(A,B,C) = [1-(1-P(C|A)) \cdot (1-P(C|B))] \cdot P(A) \cdot P(B)$$

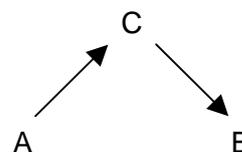
Gesetzt

$$P(a) = P(b) = 0,5 \\ P(c|a.\sim b) = P(c|b.\sim a) = 0,8 \\ P(c|\sim a.\sim b) = 0,2$$

Berechnet

$$P(a.b.c) = 0,242 \\ P(a.b.\sim c) = 0,008 \\ P(a.\sim b.c) = 0,21 \\ P(\sim a.b.c) = 0,21 \\ P(a.\sim b.\sim c) = 0,04 \\ P(\sim a.b.\sim c) = 0,04 \\ P(\sim a.\sim b.c) = 0,05 \\ P(\sim a.\sim b.\sim c) = 0,20 \\ \Sigma P(A,B,C) = 1,0$$

Kettenmodell



$$P(A,B,C) = P(B|C) \cdot P(C|A) \cdot P(A)$$

$$P(a) = P(\sim a) = 0,5 \\ P(a|c) = P(b|c) = 0,8 \\ P(a|\sim c) = P(b|\sim c) = 0,2$$

$$P(a.b.c) = 0,32 \\ P(a.b.\sim c) = 0,02 \\ P(a.\sim b.c) = 0,08 \\ P(\sim a.b.c) = 0,08 \\ P(a.\sim b.\sim c) = 0,08 \\ P(\sim a.b.\sim c) = 0,08 \\ P(\sim a.\sim b.c) = 0,02 \\ P(\sim a.\sim b.\sim c) = 0,32 \\ \Sigma P(A,B,C) = 1,0$$

Abb. 1.13: Beispiele für die Berechnung von Kausalmodellen

¹³ Da der Effekt auch eintritt, wenn keines der beiden Ursacheereignisse gegeben ist, muss dies bei der Berechnung zusätzlich mit berücksichtigt werden (vgl. Jensen, 1997, S.49).

1.3.4 Bestimmung struktureller Implikationen

Die strukturellen Implikationen eines Kausalmodells können ebenfalls auf der Grundlage des Strukturmodells und der Stärke der beteiligten Kausalrelationen errechnet werden. Wir werden dafür die oben bestimmten Wahrscheinlichkeitsverteilungen zu Grunde legen. Die folgende Abbildung zeigt für die drei Grundmodelle die Berechnung von bedingten und unbedingten Abhängigkeiten zwischen den beiden jeweils nicht direkt kausal verbundenen Ereignissen.

Die zugehörigen Wahrscheinlichkeiten werden dabei gemäß dem zweiten Axiom der Wahrscheinlichkeitsrechnung als Summe von Wahrscheinlichkeiten elementarer Ereigniskombinationen berechnet. Dieses Axiom besagt, dass die Wahrscheinlichkeit einer Vereinigung einander ausschließender Elementarereignisse gleich der Summe der Wahrscheinlichkeiten der Elementarereignisse ist, $P(a.b) = P(a) + P(b)$. In bezug auf Kausalmodelle sind die einzelnen Ereigniskombinationen die Elementarereignisse. Die Berechnung soll für $P(b|a)$ exemplarisch vorgeführt werden. Nach der Definition der bedingten Wahrscheinlichkeit ist

$$P(b|a) = P(a.b) / P(a)$$

Es gibt zwei Elementarereignisse, in denen a und b gemeinsam auftreten, a.b.c und a.b.~c, daher gilt

$$P(a.b) = P(a.b.c) + P(a.b.~c)$$

Das Ereignis a wiederum tritt in vier Elementarereignisse auf a.b.c, a.b.~c, a.~b.c und a.~b.~c.

Daraus folgt

$$P(a) = P(a.b.c) + P(a.b.~c) + P(a.~b.c) + P(a.~b.~c)$$

Damit ergibt sich die bedingte Wahrscheinlichkeit als

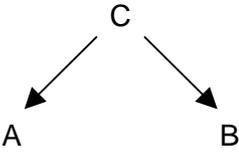
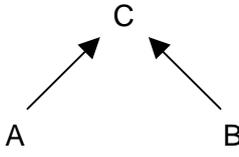
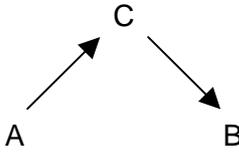
$$P(b|a) = [P(a.b.c) + P(a.b.~c)] / [P(a.b.c) + P(a.b.~c) + P(a.~b.c) + P(a.~b.~c)]$$

Setzt man nun die Werte für das Gemeinsame-Ursache Modell aus Abbildung 1.12 ein, dann resultiert

$$P(b|a) = [0,32 + 0,02] / [0,32 + 0,02 + 0,08 + 0,08] = 0,68$$

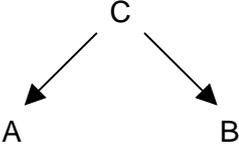
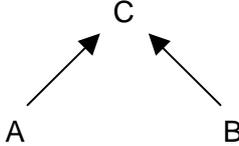
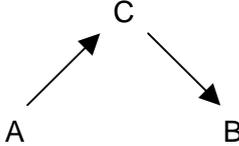
In gleicher Weise können alle anderen Wahrscheinlichkeiten berechnet werden.

Die folgende Abbildung 1.14 stellt die Bestimmung der unbedingten Abhängigkeiten zwischen den nicht direkt kausal verbundenen Ereignissen dar.

Gemeinsame-Ursache Modell	Gemeinsamer-Effekt Modell	Kettenmodell
		
$P(B A) \neq P(B)$ $P(B A,C) = P(B C)$	$P(B A) = P(B)$ $P(B A,C) \neq P(B C)$	$P(B A) \neq P(B)$ $P(B A,C) = P(B C)$
Unbedingte Abhängigkeiten		
$P(b a) = \frac{[P(a.b.c) + P(a.b.\sim c)]}{[P(a.b.c) + P(a.b.\sim c) + P(a.\sim b.c) + P(a.\sim b.\sim c)]}$ $= \frac{[0,32 + 0,02]}{[0,32 + 0,02 + 0,08 + 0,08]}$ $= 0,68$ $P(b \sim a) = 0,32$ $P(a b) = 0,68$ $P(a \sim b) = 0,32$ $P(b) = P(a.b.c) + P(a.b.\sim c) + P(\sim a.b.c) + P(\sim a.b.\sim c)$ $= 0,32 + 0,02 + 0,08 + 0,08$ $= 0,50$ $P(a) = 0,5$ $\Rightarrow P(b a) > P(b)$ $\Rightarrow P(a b) > P(a)$ $\Rightarrow \Delta P_{AB} = P(b a) - P(b) = 0,36$ $\Rightarrow \Delta P_{BA} = P(a b) - P(a) = 0,36$	$P(b a) = \frac{[P(a.b.c) + P(a.b.\sim c)]}{[P(a.b.c) + P(a.b.\sim c) + P(a.\sim b.c) + P(a.\sim b.\sim c)]}$ $= \frac{[0,242 + 0,008]}{[0,242 + 0,008 + 0,21 + 0,04]}$ $= 0,50$ $P(b \sim a) = 0,50$ $P(a b) = 0,50$ $P(a \sim b) = 0,50$ $P(b) = P(a.b.c) + P(a.b.\sim c) + P(\sim a.b.c) + P(\sim a.b.\sim c)$ $= 0,242 + 0,008 + 0,08 + 0,08$ $= 0,50$ $P(a) = 0,5$ $\Rightarrow P(b a) = P(b)$ $\Rightarrow P(a b) = P(a)$ $\Rightarrow \Delta P_{AB} = P(b a) - P(b) = 0,00$ $\Rightarrow \Delta P_{BA} = P(a b) - P(a) = 0,00$	$P(b a) = \frac{[P(a.b.c) + P(a.b.\sim c)]}{[P(a.b.c) + P(a.b.\sim c) + P(a.\sim b.c) + P(a.\sim b.\sim c)]}$ $= \frac{[0,32 + 0,02]}{[0,32 + 0,02 + 0,08 + 0,08]}$ $= 0,68$ $P(b \sim a) = 0,32$ $P(a b) = 0,68$ $P(a \sim b) = 0,32$ $P(b) = P(a.b.c) + P(a.b.\sim c) + P(\sim a.b.c) + P(\sim a.b.\sim c)$ $= 0,32 + 0,02 + 0,08 + 0,08$ $= 0,50$ $P(a) = 0,50$ $\Rightarrow P(b a) > P(b)$ $\Rightarrow P(a b) > P(a)$ $\Rightarrow \Delta P_{AB} = P(b a) - P(b) = 0,36$ $\Rightarrow \Delta P_{BA} = P(a b) - P(a) = 0,36$

*Abb. 1.14: Berechnung der strukturellen Implikationen:
Unbedingte Abhängigkeiten*

Wie zu erwarten, zeigt sich beim Gemeinsame-Ursache und beim Kettenmodell eine positive Abhängigkeit der beiden Effekte bzw. der ersten Ursache und des letzten Effektes. Dagegen sind beim Gemeinsamer-Effekt Modell die Ereignisse A und B unabhängig voneinander. Dies ergibt sowohl der Vergleich von $P(a|b)$ und $P(b)$ wie die Berechnung der Kontingenz zwischen beiden Ereignissen. Neben den unbedingten Abhängigkeiten ergeben sich für die Kausalmodelle auch spezifische bedingte Abhängigkeiten. Deren Berechnung wird in der folgenden Abbildung dargestellt.

Gemeinsame-Ursache Modell	Gemeinsamer-Effekt Modell	Kettenmodell
		
Bedingte Abhängigkeiten		
$P(b a.c) = P(a.b.c) / [P(a.b.c) + P(a.\sim b.c)]$ $= [0,32] / [0,32 + 0,08]$ $= 0,80$	$P(b a.c) = P(a.b.c) / [P(a.b.c) + P(a.\sim b.c)]$ $= [0,242] / [0,242 + 0,21]$ $= 0,54$	$P(b a.c) = P(a.b.c) / [P(a.b.c) + P(a.\sim b.c)]$ $= [0,32] / [0,32 + 0,08]$ $= 0,80$
$P(b \sim a.c) = 0,80$	$P(b \sim a.c) = 0,81$	$P(b \sim a.c) = 0,80$
$P(b a.\sim c) = 0,20$	$P(b a.\sim c) = 0,17$	$P(b a.\sim c) = 0,20$
$P(b \sim a.\sim c) = 0,20$	$P(b \sim a.\sim c) = 0,17$	$P(b \sim a.\sim c) = 0,20$
$P(b c) = [P(a.b.c) + P(\sim a.b.c)] / [P(a.b.c) + P(\sim a.b.c) + P(a.\sim b.c) + P(\sim a.\sim b.c)]$ $= [0,32 + 0,08] / [0,32 + 0,08 + 0,08 + 0,02]$ $= 0,80$	$P(b c) = [P(a.b.c) + P(\sim a.b.c)] / [P(a.b.c) + P(\sim a.b.c) + P(a.\sim b.c) + P(\sim a.\sim b.c)]$ $= [0,242 + 0,21] / [0,242 + 0,21 + 0,21 + 0,05]$ $= 0,64$	$P(b c) = [P(a.b.c) + P(\sim a.b.c)] / [P(a.b.c) + P(\sim a.b.c) + P(a.\sim b.c) + P(\sim a.\sim b.c)]$ $= [0,32 + 0,08] / [0,32 + 0,08 + 0,08 + 0,02]$ $= 0,80$
$\Rightarrow P(b a.c) = P(b c)$	$\Rightarrow P(b a.c) < P(b c)$	$\Rightarrow P(b a.c) = P(b c)$
$\Rightarrow \Delta P_{AB,C} = P(b a.c) - P(b \sim a.c)$ $= 0,00^{14}$	$\Rightarrow \Delta P_{AB,C} = P(b a.c) - P(b \sim a.c)$ $= -0,27$	$\Rightarrow \Delta P_{AB,C} = P(b a.c) - P(b \sim a.c)$ $= 0,00$
$\Rightarrow \Delta P_{AB,\sim C} = P(b a.\sim c) - P(b \sim a.\sim c)$ $= 0,00$	$\Rightarrow \Delta P_{AB,\sim C} = P(b a.\sim c) - P(b \sim a.\sim c)$ $= 0,00$	$\Rightarrow \Delta P_{AB,\sim C} = P(b a.\sim c) - P(b \sim a.\sim c)$ $= 0,00$

*Abb. 1.15: Berechnung der strukturellen Implikationen:
Bedingte Abhängigkeiten*

Für die bedingten Abhängigkeiten zeigt sich ebenfalls das zu erwartende Bild. Beim Gemeinsame-Ursache Modell und beim Kettenmodell werden die Ereignisse A und B durch das Konstanthalten der gemeinsamen Ursache bzw. des intermittierenden Ereignisses unabhängig voneinander. Ganz anders beim Gemeinsamer-Effekt Modell. Liegt der Effekt vor, dann sind die beiden Ursachen abhängig voneinander und zwar negativ. Liegen der Effekt und eine der Ursache vor, dann sinkt die Wahrscheinlichkeit der anderen Ursache.

¹⁴ Da die bedingten Wahrscheinlichkeiten für beide Kausalzusammenhänge gleich sind, sind die Kontingenzen zwischen den beiden Effekten symmetrisch. Aus Gründen einer besseren Übersichtlichkeit wurde daher auf die Angabe der anderen Kontingenz verzichtet. Dies gilt entsprechend auch für die anderen Modelle.

1.4 Zusammenfassung

Ziel dieses Einführungskapitels war es, qualitative, formale und quantitative Aspekte von Kausalmodellen aufzuzeigen. Dabei stellte sich die Struktur der Modelle als deren wichtigste Eigenschaft heraus. Verschiedene Modelle unterscheiden sich vor allem durch ihre Struktur voneinander. Diese Strukturen implizieren gewisse bedingte und unbedingte Abhängigkeiten auch für nicht direkt kausal verbundene Ereignisse. Insbesondere das Gemeinsamer-Effekt Modell weist strukturelle Implikationen auf, die kein anderes Modell hat. Dadurch, dass Kausalmodelle nur über ihre Struktur bestimmt werden, können sie auf jegliche inhaltliche Domäne angewandt werden. Über ihre Struktur können beobachtete Kovariationen sowohl zwischen kausal verbundenen wie zwischen nicht direkt kausal verbundenen Ereignissen erklärt werden. So erklärt ein Gemeinsame-Ursache Modell nicht nur, weshalb bei Gewittern das Barometer fällt und es regnet, sondern auch, warum der Stand des Barometers und Regen zusammenhängen, obwohl weder der Stand des Barometers den Regen noch der Regen den niedrigen Barometerstand auslöst. Strukturgleichungsmodelle oder Bayesianische Modelle erlauben es, diese strukturellen Eigenschaften formal abzubilden. Dadurch werden Kausalmodelle nicht nur exakt beschreibbar, sie können auch als Computermodelle implementiert werden. Dies wiederum ermöglicht es, Modelle sehr komplexer Sachverhalte zu bauen und diese für Erklärungen wie für Vorhersagen zu nutzen. So zum Beispiel für die Vorhersage des Wetters (Edwards, 1998) oder die Höhe des Ölpreises (Abramson & Finizza, 1995).

In diesem ersten Kapitel sind wir nicht auf psychologische Aspekte von Kausalmodellen eingegangen. Kausalmodelle wurden von ihrer normativen Seite betrachtet. Diese formalen und normativen Aspekte liefern einen wichtigen Vergleichsmaßstab, wenn wir uns im restlichen Verlauf dieser Arbeit mit Kausalmodellen als mentalen Repräsentationen von Ursache-Wirkungszusammenhängen beschäftigen. Die zentrale Frage wird dabei sein, ob Personen sensitiv für die normativen Unterschiede zwischen verschiedenen Kausalmodellen sind. Erkennen sie die Bedeutung der Strukturen? Wissen sie, worin sich verschiedene Modelle unterscheiden und welche Implikationen sich aus diesen ergeben? Die folgenden Kapitel werden versuchen, erste Antworten auf diese Fragen zu geben.

2. ANWENDUNGEN VON KAUSALMODELLEN

Im ersten Kapitel wurde die Frage beantwortet, worin sich verschiedene Kausalmodelle unterscheiden. Das zweite Kapitel versucht erste Antworten auf die Frage zu geben, ob Personen für die aufgezeigten Unterschiede sensitiv sind. Dass Personen häufig Kausalmodelle zur Veranschaulichung einsetzen, ist offensichtlich. Graphen zur Darstellung von Abläufen und komplexen Ursache-Wirkungszusammenhängen sind in der Wissenschaft, in der Wirtschaft und im Alltag weit verbreitet (vgl. Oestermeier & Hesse, 2000). Nutzen Personen aber auch die strukturellen Eigenschaften dieser Modelle? Erkennen sie die Unterschiede zwischen den Strukturen und die daraus resultierenden Implikationen?

Kausalmodelle sind für eine Vielzahl von Aufgaben von Bedeutung. Das zweite Kapitel gibt einen Überblick. Als Erstes wird es dabei um den Einfluss von Kausalmodellen auf das Erlernen von Kausalzusammenhängen gehen. Im einfachsten Fall ist der Zusammenhang zwischen zwei Ereignissen zu bestimmen. Bereits hierfür ist wichtig zu wissen, welches der beiden Ereignisse die Ursache ist. Liegt eine Vielzahl von Ereignissen vor, so sind je nach Kausalmodell unterschiedliche statistische Zusammenhänge für die Stärke der zu Grunde liegenden Kausalzusammenhänge indikativ. Im zweiten Teilkapitel werden wir uns dann Prognosen und Diagnosen zuwenden. Die Aufgabe bei Prognosen ist es, aus dem Vorliegen einer bestimmten Ursachenkonstellation auf das Eintreten von Effekten zu schließen. Diagnosen verfolgen das entgegengesetzte Ziel. Aus dem Vorliegen von Effekten soll die Ursachenkonstellation rekonstruiert werden. Beide Prozesse können dabei auf Gemeinsamer-Effekt Modellen wie Gemeinsame-Ursache Modellen beruhen. Kausalmodelle sind aber in der Lage noch mehr zu leisten. Dank ihrer strukturellen Implikationen können auch nicht direkt kausal verbundene Ereignisse vorhergesagt werden. Das dritte Teilkapitel befasst sich mit der Überprüfung von Kausalhypothesen als einem weiteren Anwendungsfeld. Hierbei wird es zunächst um die Überprüfung von kausalen Zusammenhangshypothesen gehen. Diese Aufgabe ist durch die Bestimmung des normativ richtigen Zusammenhangs einfach zu lösen. Dagegen verlangt die Überprüfung von Strukturhypothesen ein anderes Vorgehen. Um herauszufinden, ob ein vermutetes Kausalmodell tatsächlich gegeben ist oder nicht, müssen neben den Kausalzusammenhängen auch die strukturellen Implikationen überprüft werden. Nur wenn diese durch die Daten bestätigt werden, kann das Modell zutreffen. Im vierten Teilkapitel werden wir uns mit der Generierung von Kausalhypothesen beschäftigen. Eine Aufgabe ist es, eine Vielzahl beobachteter Kovariationen auf eine begrenzte Zahl von Ursachefaktoren zurückzuführen. Kausalmodelle liefern hierfür wichtige Constraints. Sie geben vor, wie viele Ursachen benötigt werden. Eine andere Aufgabe ist es, ein Kausalmodell zu induzieren. Ein möglicher

Weg ist es, das Modell aus den statistischen Zusammenhängen zu erschließen. Hierfür müssen alle bedingten und unbedingten Abhängigkeiten zwischen den beteiligten Ereignissen ermittelt werden. Durch einen Vergleich mit den strukturellen Implikationen verschiedener Modelle kann dann das Richtige herausgefunden werden. Ein zweiter Weg ist es, durch Experimente herauszufinden, welche Ereignisse welche anderen Ereignisse beeinflussen. Der einfachste und universell einsetzbare Weg ist es aber, mehrere einzelne Kausalzusammenhänge zu einem Modell zu verbinden. Wenn die Integration erfolgreich durchgeführt wird, dann lassen sich auch über die strukturellen Implikationen des Modells die Zusammenhänge zwischen den nicht kausal verbundenen Ereignissen vorhersagen. Die unten stehende Abbildung zeigt nochmals die Struktur des Kapitels im Überblick.

2. ANWENDUNGEN VON KAUSALMODELLEN							
2.1	Lernen von Kausalzusammenhängen	2.2	Prognosen und Diagnosen	2.3	Prüfen von Kausalhypothesen	2.4	Generieren von Kausalhypothesen
2.1.1	Lernen einzelner Kausalzusammenhänge	2.2.1	Prognose von Effekten	2.3.1	Prüfen von Zusammenhangshypothesen	2.4.1	Generieren von kausalen Zusammenhangshypothesen
2.1.2	Lernen von Kausalmodellen	2.2.2	Diagnose von Ursachen	2.3.2	Prüfen von Kausalzusammenhängen innerhalb eines Kausalmodells	2.4.2	Generieren von kausalen Strukturhypothesen
		2.2.3	Vorhersage nicht direkt kausal verbundener Ereignisse	2.3.3	Prüfen von kausalen Strukturhypothesen		
2.5	Zusammenfassung						

Abb. 2.1: Struktur Kapitel 2

Innerhalb der einzelnen Teilkapitel werden zunächst die Aufgabenstellung und das je nach Kausalmodell normativ richtige Vorgehen erläutert. Danach werden die Befunde relevanter Studien zum diesem Thema besprochen. Aus den noch offenen Fragen werden dann jeweils die Fragestellungen und die grundlegenden Versuchsparadigmen der im Rahmen dieser Arbeit durchgeführten Studien entwickelt.

2.1 Lernen von Kausalzusammenhängen

Beim Lernen von Kausalzusammenhängen muss aus einer Vielzahl von Lernerfahrungen auf das Vorhandensein eines Kausalzusammenhangs geschlossen werden. Dabei können die statistischen Zusammenhänge zwischen den Ereignissen als Indikatoren für den Kausaleinfluss herangezogen werden. Wie aber bereits im ersten Kapitel dargelegt, sind nur bestimmte statistische Zusammenhänge wirklich indikativ für Kausalität. Welche dies sind, hängt von dem zu Grunde liegenden Kausalmodell ab. Das folgende Teilkapitel beschäftigt sich mit dem Einfluss von Kausalmodellen auf das Lernen einzelner Kausalzusammenhänge (2.1.1) und auf das gleichzeitige Lernen mehrerer Kausalzusammenhänge (2.1.2).

2.1.1 Lernen einzelner Kausalzusammenhänge

Bereits beim Lernen einzelner Kausalzusammenhänge zwischen zwei Ereignissen ist die Struktur des angenommenen Kausalmodells entscheidend. Struktur bedeutet in diesem Fall die Richtung des Kausalzusammenhangs. Verursacht das Ereignis A das Ereignis B oder ruft Ereignis B Ereignis A hervor? Beeinflussen beispielsweise die Einstellungen das Verhalten oder das Verhalten die Einstellungen? Wie die Diskussion der quantitativen Aspekte von Kausalmodellen gezeigt hat, ist die Kontingenz der beiden Ereignisse indikativ für die Stärke des Kausalzusammenhangs. Die Kontingenz ist aber abhängig von der Richtung des angenommenen Kausalzusammenhangs. Ist A die Ursache von B, so gibt $\Delta P = P(b|a) - P(b|\sim a)$ die Wirksamkeit an; ist A der Effekt von B, dann muss $\Delta P = P(a|b) - P(a|\sim b)$ bestimmt werden. Der gleiche statistische Zusammenhang von A und B kann also zwei unterschiedlich starke Kausalzusammenhänge bedeuten.¹⁵ Die Abbildung auf der folgenden Seite zeigt ein Beispiel für einen solchen Unterschied.

¹⁵ Die gleiche Überlegung trifft auch für alle anderen Maße von Kausalität zu, da diese auf der Kontingenz beruhen. D.h., auch die kausale Power und die ‚Probability of necessity and sufficiency‘ sind sensitiv für die Richtung des Kausalzusammenhangs.

	Ereignis B	Kein Ereignis B
Ereignis A	48	32
Kein Ereignis A	0	8
Angenommene Kausalrichtung	A → B	B → A
Kontingenzen	$\Delta P_{AB} = P(b a) - P(b \sim a)$ = 32/80 - 0/8 = 0,80	$\Delta P_{BA} = P(a b) - P(a \sim b)$ = 48/48 - 32/40 = 0,20

Abb. 2.2: Zahlenbeispiel für den Einfluss der Richtung des Kausalzusammenhangs auf die Bestimmung der Stärke des Kausalzusammenhangs

Wenn anzunehmen ist, dass Ereignis A Ereignis B verursacht, dann resultiert ein sehr starker Kausalzusammenhang von $\Delta P_{AB} = 0,80$, wenn aber die Vermutung ist, dass Ereignis B Ereignis A hervorruft, dann sprechen die Daten für einen lediglich schwachen Zusammenhang von $\Delta P_{BA} = 0,20$.

Empirische Befunde

Eine Vielzahl von Experimenten beschäftigte sich mit der Frage, wie Probanden¹⁶ beim Erlernen von Kausalzusammenhängen vorgehen bzw. ob sie sensitiv für die Kontingenzen zwischen der Ursache und dem Effekt sind. Die Aufgabe der Probanden war es, den Zusammenhang zwischen zwei Ereignissen zu bestimmen. So wurden sie zum Beispiel gebeten herauszufinden, ob eine Impfung von Wolken mit einer chemischen Substanz zu Regen führt (Jenkins & Ward, 1965; Ward & Jenkins, 1965). Als Daten erhielten die Probanden Informationen über das Auftreten der beiden Ereignisse. So können beide Ereignisse entweder gemeinsam auftreten (u.e), gemeinsam abwesend sein ($\sim u.\sim e$), die Ursache kann allein auftreten ($\sim u.e$) oder der Effekt kann ohne vorherige Ursache vorkommen (e. $\sim u$). Die untenstehende Abbildung zeigt die hierfür übliche Darstellung in Form einer Vier-Felder-Tafel.

	Effekt	\sim Effekt
Ursache	Zelle a u.e	Zelle b u. $\sim e$
\sim Ursache	Zelle c $\sim u.e$	Zelle d $\sim u.\sim e$

Abb. 2.3: Vier-Felder Tafel der möglichen Ereigniskombinationen in einem Experiment zum Kausallernen

¹⁶ Mit Probanden sind natürlich stets Männer wie Frauen gemeint.

Die Kontingenz zwischen der Ursache und dem Effekt ergibt sich aus den Häufigkeiten der jeweiligen Ereigniskombinationen über

$$\Delta P = [f(e.u) / f(e.u) + f(\sim e.u)] - [f(e.\sim u) / f(e.\sim u) + f(\sim e.\sim u)] \text{ (Kontingenz)}$$

oder bezogen auf die Zellen als

$$\Delta P = a / (a+b) - c / (c+d) \text{ (Kontingenz)}$$

Das bedeutet, dass alle vier Ereigniskombinationen gleichermaßen relevant für die Bestimmung der Kontingenz und alle auf ihr aufbauenden Maße für Kausalität (z.B. Power) sind. In älteren Untersuchungen zeigte sich aber, dass nur eine Minderheit der Probanden den Zusammenhang über die Kontingenz bestimmte (Jenkins & Ward, 1965; Ward & Jenkins, 1965). Die Mehrheit der Probanden wendete eine Strategie an, die als Summe der Diagonalen (Inhelder & Piaget, 1958) oder ΔD bezeichnet wird. Der Zusammenhang der Ereignisse wird danach bestimmt über

$$\Delta D = (a+d) - (b+c) \text{ (Summe der Diagonalen)}$$

In verschiedenen Untersuchungen dominierte dieses Vorgehen (Shaklee & Tucker, 1980).¹⁷ Die Probanden werteten das gemeinsame Auftreten und das gemeinsame Ausbleiben der Ursache und ihres Effektes als positive Belege für einen Zusammenhang. Davon subtrahierten sie als Gegenbelege, wenn die Ursache nicht vom Effekt gefolgt wurde oder der Effekt ohne Ursache auftrat.

Regressionsanalysen über die Einschätzungen des Kausalzusammenhangs mit den vier Ereigniskombinationen als Prädiktoren ergaben, dass die einzelnen Ereigniskombinationen bei der Urteilsbildung unterschiedlich stark gewichtet wurden. So wurde das gemeinsame Auftreten von Ursache und Effekt als stark positive Evidenz gewertet. Traten die Ursache oder der Effekt alleine auf, so wurde dies als Gegenbeleg angesehen und negativ gewichtet. Das gemeinsame Ausbleiben von Ursache und Effekt wurde entweder als neutral oder als leicht positive Evidenz gewertet (Kao & Wasserman, 1993; Schustack & Sternberg, 1980). Über diese Regressionsmodelle ließ sich ein Großteil der Varianz (bis zu 90%) in den Urteilen aufklären (Schustack & Sternberg, 1980; White, 2000).

Ein normativ richtiges Vorgehen nach der ΔP -Regel wurde unter gewissen Bedingungen gefunden (Alloy & Abramson, 1979; Ward & Jenkins, 1965). Die Probanden verwendeten die Kontingenzen v.a. dann, wenn die Ereignisse neutral und die Kontingenzen positiv waren. Nullkontingenzen wurden erkannt, wenn die Ereignisse nicht zu häufig gemeinsam auftraten (siehe Alloy & Tabachnik, 1984, für einen Überblick).

¹⁷ Werden die Anforderungen an die kognitiven Kapazitäten erhöht, dann wenden die Probanden häufig eine noch einfachere Strategie an, sie subtrahieren die Anzahl der Fälle in der A-Zelle von denen in der B-Zelle (Shaklee & Mims, 1982).

Die unterschiedlichen Vorgehensweisen führen häufig zu denselben Urteilen. So korrelieren die Kausalurteile, die auf Basis der ΔD -Regel gefällt werden, sehr hoch mit denen, die auf Kontingenzen beruhen. Wenn die Ursache in etwa gleich häufig eintritt wie ausbleibt, konvergieren beide Maße (Shaklee & Tucker, 1980). In Simulationsstudien konnte gezeigt werden, dass die Ergebnisse verschiedener Strategien hoch miteinander korrelieren (McKenzie, 1994). Die ΔD -Regel kann daher als eine gute vereinfachende Heuristik für die Bestimmung der Kontingenz angesehen werden.

Forschungsarbeiten zum instrumentellen Lernen haben sich ebenfalls mit der Einschätzung von Kausalzusammenhängen beschäftigt. Als Ursachen fungierten in diesen Untersuchungen instrumentelle Handlungen, die bestimmte Effekte hervorriefen. Ein Beispiel war das Drücken einer Taste, was zu einem Aufleuchten einer Lampe führte. Die Aufgabe war es einzuschätzen, wie stark über das Drücken der Taste die Lampe beeinflusst werden kann. Das Ergebnis einer Vielzahl von Experimenten war, dass die geschätzten Zusammenhänge sehr genau den in den Daten vorgegebenen Kontingenzen entsprachen. (Shanks, 1993; Shanks & Dickinson, 1987; Wasserman et al. 1993).

Andere Untersuchungen haben sich in letzter Zeit mit der Frage beschäftigt, ob Personen Kontingenzen oder kausale Power als Indikatoren für die Wirksamkeit einer Ursache heranziehen (Buehner & Cheng, 1997; Lien & Cheng, 2000; Wu & Cheng, 1999). Hierzu wurden Probanden gebeten, die Stärke eines Kausalzusammenhangs einzuschätzen. In verschiedenen Experimenten wurden sowohl die Powerwerte als auch die Kontingenzen zwischen der Ursache und dem Effekt manipuliert. Die Befunde zeigen kein einheitliches Bild. Wurden die Kontingenzen konstant gehalten, aber die Powerwerte variiert, dann stiegen die Urteile parallel zu den Powerwerten an. Wurden aber die Powerwerte konstant gehalten und die Kontingenzen variiert, dann hatte dies ebenfalls eine Veränderung der Schätzurteile zur Folge. Wie diese Befunde zu erklären sind, darüber wird derzeit noch heftig debattiert (Lober & Shanks, 2000). Festzuhalten ist, dass Personen sowohl Kontingenzen als auch kausale Power als Indikatoren für die Bestimmung der Stärke eines Kausalzusammenhangs heranziehen können.

In verschiedenen Studien wurde der Einfluss der Datendarbietung auf die Kausalitätsschätzungen untersucht. So zeigte sich, dass eine zusammenfassende Darstellung der Daten in Tabellenform akkuratere Einschätzungen der Kontingenz zur Folge hat als eine Darbietung in Form von Einzelfällen. So nutzten in der klassischen Studie von Ward & Jenkins (1965) nur 17% der Probanden bei einer Trial-by-Trial Lernaufgabe die Kontingenz, aber 75% bei einer summarischen Darstellung der Daten. Spätere Studien

konnten aber belegen, dass auch beim Trial-by-Trial Lernen genaue Einschätzungen gemacht werden (Shanks, 1993; Shanks et al., 1996; Spellman, 1996a,b; Wasserman et al., 1993).

Zusätzlich wurde ein Effekt der Abfolge der Lerntrials nachgewiesen (Dennis & Ahn, *subm.*; Hogarth & Einhorn, 1992). Die Kontingenz, welche die Ereignisse zu Beginn einer Datenserie haben, bestimmt das Urteil stärker als die Kontingenz über alle Lerntrials hinweg. Zum Beispiel wurde Probanden ein Datensatz vorgelegt, bei dem die gezeigten Ereignisse zu Beginn des Lernens eine positive und zu Ende eine negative Kontingenz aufwiesen. Insgesamt waren die Ereignisse unkorreliert. Das Ergebnis war, dass der Kausalzusammenhang trotz dieser Nullkontingenz positiv eingeschätzt wurde (Dennis & Ahn, *subm.*). Dieser Primacy-Effekt zeigte sich auch in anderen Experimenten (Ahn & Dennis, 2000).

Untersuchungen zum wissenschaftlichen Denken beschäftigen sich ebenfalls mit dem Erwerb von Wissen über Kausalzusammenhänge.¹⁸ Hierbei wurde u.a. untersucht, ob Personen in der Lage sind, verschiedene Formen von Kausalzusammenhängen zu entdecken (Kuhn et al., 1995). Die Probanden hatten die Aufgabe, den Einfluss von fünf verschiedenen Ursachen auf einen Effekt einzuschätzen. Eine der Aufgaben war es beispielsweise, herauszufinden, wovon die Popularität eines Fernsehprogramms abhing. Andere Aufgaben betrafen physikalische Probleme, z.B. Faktoren, von denen die Geschwindigkeit eines Spielbootes in einem Wassertank abhängt (vgl. auch Schauble, 1996). Von den fünf Kausalzusammenhängen war einer klar positiv, zwei Faktoren hatten keinen Einfluss auf den Effekt, die vierte Ursache hatte nur in Interaktion mit der ersten Ursache eine Wirkung und der fünfte Zusammenhang war kurvilinear. Das Ergebnis war, dass die Probanden nur den positiven und die beiden Nullzusammenhänge richtig ermittelten. Die Wechselwirkung und der kurvilineare Zusammenhang wurden sehr selten erkannt (Kuhn et al., 1995).

Weitere Untersuchungen haben sich mit der Rolle des Vorwissens beim Bestimmen einfacher Kausalzusammenhänge beschäftigt. Die Frage war, ob das Wissen über mögliche Mechanismen, welche die Ursache und den Effekt verbinden können, die Einschätzung beeinflusst (Koslowski, 1996). Eine Aufgabe war es zum Beispiel, herauszufinden, ob durch bestimmte Maßnahmen die Verweildauer von Patienten in Krankenhäusern verkürzt werden

¹⁸ Viele dieser Untersuchungen haben entwicklungspsychologische Fragestellungen. Auf diese Ergebnisse kann aus Platzgründen nicht eingegangen werden. Die zitierten Befunde beziehen sich ausschließlich auf die Erwachsenenvergleichsgruppe.

kann. Als mögliche Ursache wurde entweder die Veränderung der Vorhangfarbe (kein Vorwissen über mögliche Mechanismen) oder die Verlängerung der Besuchszeiten (mit Vorwissen über mögliche Mechanismen) vorgegeben. Die Kontingenz zwischen der Ursache und dem Effekt wurde in beiden Bedingungen konstant hoch gehalten. Das Ergebnis in einer Vielzahl von Experimenten mit unterschiedlichen Aufgabenstellungen aus dem sozialen, technischen und physikalischen Bereich war, dass trotz identischer Kontingenz der Kausalzusammenhang deutlich schwächer eingeschätzt wurde, wenn kein plausibler Mechanismus bekannt war, der die Wirkung vermitteln konnte (Koslowski, 1996). Dieser Effekt trat sogar bei deterministischen Zusammenhängen auf. Ebenso zeigte es sich, dass die Einschätzungen zurückgehen, wenn nachträglich das Vorliegen eines plausiblen Wirkungsmechanismus ausgeschlossen wird (vgl. Koslowski, 1996 für einen Überblick zu Arbeiten, die sich mit der Rolle des Vorwissens beschäftigen).

Fassen wir die obigen Befunde zusammen, dann ergibt sich folgendes Bild: Probanden sind sensitiv für die Kontingenzen zwischen Ereignissen und nutzen den statistischen Zusammenhang für die Beurteilung des kausalen Zusammenhangs. Bei der Bestimmung der Stärke des Kausalzusammenhangs gehen die Probanden in vielen Fällen normativ vor. Sie ermitteln also die Kontingenz bzw. die Power zwischen der Ursache und ihrer Wirkung. Abweichungen von diesem normativen Vorgehen zeigen sich vor allem unter zwei Bedingungen. Wenn hohe Anforderungen an die Verarbeitungskapazität gestellt werden, beispielsweise durch eine Trial-by-Trial Darbietung der Daten, dann setzen Probanden häufig vereinfachende Heuristiken ein. Wenn die Ereignisse, zwischen denen ein Zusammenhang bestimmt werden soll, bekannt sind, dann führt dieses Vorwissen zu einer Verzerrung des Urteils. Ein normatives Vorgehen mit hoch präzisen Einschätzungen der Kontingenz zeigt sich nach instrumentellem Lernen, d.h. nach eigenen Handlungen, und bei neutralen Inhalten mit einer übersichtlichen, summarischen Darstellung der Daten.

2.1.2 Lernen von Kausalmodellen

Auch beim gleichzeitigen Erlernen mehrerer Kausalzusammenhänge ist das Kausalmodell, welches die Zusammenhänge bilden, entscheidend. Wie wir bei der Diskussion der quantitativen Aspekte (vgl. Kapitel 1.3) gesehen haben, muss bei der Bestimmung der Stärke eines Kausalzusammenhangs das Kausalmodell berücksichtigt werden, in das dieser Zusammenhang eingebettet ist. Ist der Zusammenhang Teil eines Gemeinsame-Ursache Modells, dann erfasst die unbedingte Kontingenz bzw. die unbedingte Power die Stärke des Kausalzusammenhangs. Bilden die Zusammenhänge aber ein Gemeinsamer-Effekt Modell, dann droht eine Konfundierung der Wirkung der beiden

Ursachen. Der Einfluss der jeweils anderen Ursachen muss also wechselseitig kontrolliert werden. In diesem Fall geben die auf die Abwesenheit der jeweils anderen Ursache bedingten Kontingenzen bzw. die bedingten Powerwerte den Kausalzusammenhang wieder.

Empirische Befunde

Die Fragestellung einer Reihe von Experimenten war, ob Kausalmodelle einen Einfluss auf das Lernen auch bei klassischen Lernparadigmen haben (Waldmann, 1994, 2000, in press; Waldmann & Holyoak, 1992). Eines war das sogenannte Overshadowing Paradigma. In diesem Paradigma erhielten die Probanden die Aufgabe, die Zusammenhänge zwischen zwei Prädiktorereignissen (Cues) und einem Kriteriumsergebnis (Outcome) zu erlernen. Die präsentierten Daten waren so gestaltet, dass die beiden Prädiktorereignisse entweder beide gemeinsam auftraten oder beide ausblieben. Traten sie auf, dann lag auch das Kriteriumsergebnis vor, ansonsten nicht. Die beiden Prädiktorereignisse hatten also beide eine unbedingte Kontingenz von $\Delta P = 1,0$ auf das Kriteriumsergebnis. Das klassische Overshadowing Phänomen ist, dass trotz des perfekt deterministischen Zusammenhangs die Probanden beide Prädiktorereignisse als nur mittelmäßig prädiktiv für das Kriteriumsergebnis beurteilen.¹⁹ In mehreren Experimenten wurden nun die Annahmen über das hinter den Zusammenhängen stehende Kausalmodell variiert. Die Abbildung auf der folgenden Seite zeigt das Versuchsdesign.

¹⁹ Dieser Befund kann mit Theorien zum assoziativen Lernen, wie dem Rescorla-Wagner Modell erklärt werden (Rescorla & Wagner, 1972, vgl. auch Shanks & Dickinson, 1987; Wasserman et al., 1996). Nach diesem Modell wird der Wert des Outcome durch die Cues vorhergesagt. Der vorhergesagte Wert des Outcome wird dabei als Summe der mit ihrer Assoziationsstärke (v) gewichteten Cuewerte (u) bestimmt ($\sum V \cdot U$). Weicht dieser vorhergesagte Wert vom tatsächlich beobachteten Wert λ ab, so werden die Assoziationsstärken der bei dem jeweiligen Versuchsdurchgang vorhandenen Cues entsprechend erhöht oder erniedrigt ($\Delta V = \alpha\beta (\lambda - (\sum V \cdot U))$). Der Parameter Alpha berücksichtigt zusätzlich die Salienz der Cues und des Outcome. Der Parameter Beta steht für die Lernrate, d.h., er gibt an, wie stark sich die Assoziationsgewichte nach einem trial ändern. In dem vorgestellten Experiment haben die Cues und der Outcome nur die Werte 0 und 1. Beim Lernen steigt durch die wiederholte Paarung beider Cues mit dem Outcome deren Assoziationsstärke zunächst rasch an. Da aber der vorhergesagte Wert des Outcome die Summe beider mit ihrer Assoziationsstärke gewichteten Cues ist, vermag deren Assoziationsstärke nicht über einen mittleren Wert zu steigen. Sind beide Cues vorhanden und das Assoziationsgewicht liegt bei 0,5, dann ist $\sum V \cdot U = 1,0 = \lambda$. Da keine Abweichung mehr gegeben ist, können die Assoziationsgewichte nicht weiter steigen (vgl. auch Waldmann, in press, für eine ausführliche Diskussion dieses Falles).

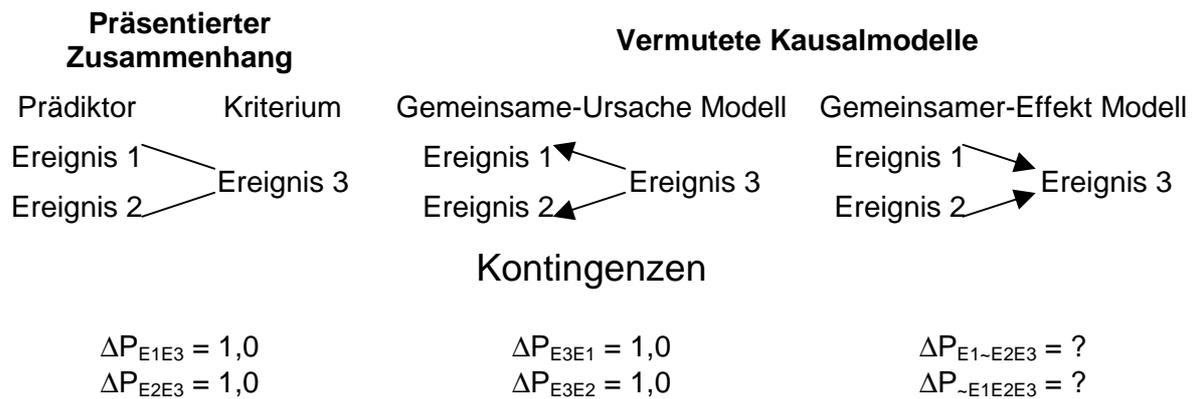


Abb. 2.4: Overshadowing Paradigma beim Kausallernen

Bei Annahme eines Gemeinsamer-Effekt Modells (rechte Spalte) waren die Prädiktorereignisse die Ursachen eines gemeinsamen Effektes. In diesem Fall sollten die Probanden die bedingten Kontingenzen als Maß für die Stärke des Kausalzusammenhangs verwenden und darauf ihr Urteil über die Vorhersagekraft des jeweiligen Ereignisses stützen. Die bedingten Kontingenzen waren aber bei den gegebenen Daten nicht zu bestimmen, da beide Ursachen stets gemeinsam auftraten. Deshalb war zu erwarten, dass die Probanden sich unsicher über die Wirksamkeit und damit auch über die Prädiktivität der einzelnen Ursachen sein sollten. Ihre Urteile sollten daher nicht hoch sein. Ganz anders aber bei Annahme eines Gemeinsame-Ursache Modells (mittlere Spalte). In diesem Fall waren die beiden Prädiktorereignisse die Effekte einer gemeinsamen Ursache. Damit wurden aber die unbedingten Kontingenzen zum richtigen Indikator für Kausalität. Diese ließen sich aufgrund der Daten leicht als $\Delta P = 1,0$ bestimmen. Die eine Ursache führte also deterministisch zu ihren beiden Effekten. Da sie hinreichend und notwendig für ihre Effekte war, konnte auch von jedem einzelnen Effekt direkt auf die Ursache zurückgeschlossen werden. D.h., beide Prädiktorereignisse sollten bei Annahme einer Gemeinsame-Ursache Struktur als hoch prädiktiv für das Kriteriumereignis beurteilt werden. Es sollte sich kein Overshadowing zeigen. Das Ergebnis des Experimentes war, dass nur bei einem Gemeinsamer-Effekt Modell eine gegenseitige „Überschattung“ auftrat, diese aber bei einem Gemeinsame-Ursache Modell ausblieb (Waldmann, in press).²⁰

Gleiche Befunde zeigten sich auch in einem Blocking Paradigma (Waldmann et al., 1992; Waldmann, 1994, 2000). Bei diesem lernten die Probanden in einer ersten Lernphase einen deterministischen Zusammenhang zwischen zwei Ereignissen. In einer zweiten

²⁰ Diese Differenz zwischen den beiden Kausalmodellen können Assoziationstheorien nicht erklären, da diese nicht zwischen der Abfolge der Ereignisse und den dahinter stehenden Kausalrelationen unterscheiden. Innerhalb dieses Experimentes waren in allen Bedingungen Cues und Outcome

Lernphase kam ein weiteres Prädiktorereignis hinzu, dass stets mit dem ersten Prädiktorereignis gemeinsam auftrat. Beide Prädiktoren sagten in dieser Phase perfekt das Kriteriumsereignis vorher. Das klassische Phänomen (Kamin, 1969) war, dass das zweite Prädiktorereignis, obwohl es in der zweiten Lernphase perfekt mit dem Kriteriumsereignis kovarierte, als nicht prädiktiv beurteilt wurde. Es wurde „blockiert“. Das Prädiktorereignis aus der ersten Phase wurde dagegen in beiden Phasen als hoch prädiktiv eingeschätzt. Nun wurden in mehreren Experimenten die Annahmen über das zu Grunde liegende Kausalmodell variiert. Die Prädiktorereignisse wurden entweder als Effekte einer gemeinsamen Ursache oder als Ursachen eines gemeinsamen Effektes vorgestellt. Die unten stehende Abbildung zeigt das zugehörige Versuchsdesign.

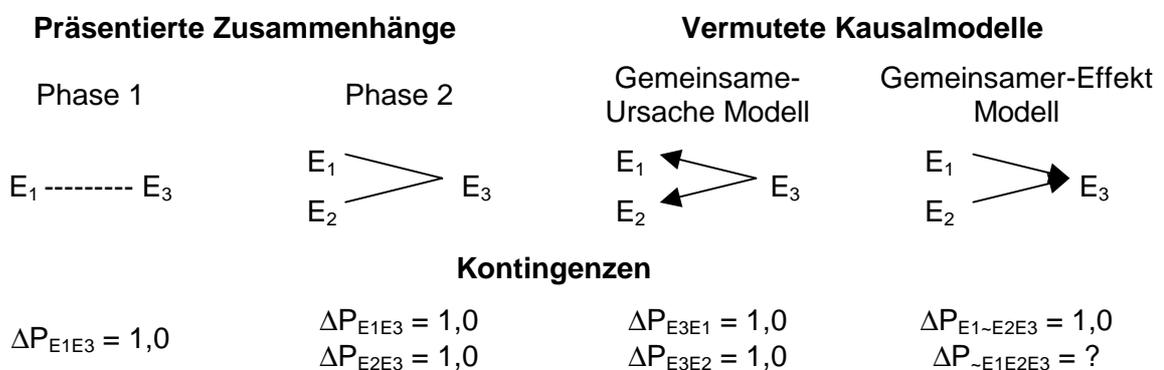


Abb. 2.5: Blocking Paradigma beim Kausallernen

Wurde das hinter den Ereignissen vermutete Kausalmodell manipuliert, dann resultierte Blocking nur bei Annahme eines Gemeinsamer-Effekt Modells. In diesem Fall war, wie zuvor beim Overshadowing Paradigma, die gesuchte bedingte Kontingenz der zweiten Ursache auf den Effekt nicht bestimmbar. Gingen die Probanden aber von einer Gemeinsame-Ursache Struktur aus, dann verschwand die Blockierung. Beide Effekte wurden als hoch prädiktiv für ihre Ursache angesehen (Waldmann 2000; Waldmann & Holyoak, 1992).²¹ Zusätzlich konnte gezeigt werden, dass diese Befunde nicht davon abhingen, ob die Probanden die Zusammenhänge prädiktiv (von der Ursache hin zur Wirkung) oder diagnostisch (ausgehend von der Wirkung hin zur Ursache) gelernt hatten (Waldmann, 2000). Als wichtig erwies sich aber, ob die Struktur des Kausalmodells den

identisch, weshalb sich nach dem Rescorla-Wagner Modell in beiden Fällen Overshadowing zeigen sollte.

²¹ Assoziationstheorien können wiederum nur den Blocking-Effekt erklären. Da der in Phase 1 gelernte Cue den Outcome perfekt vorhersagt, tritt in Phase zwei keinerlei Differenz zwischen dem vorhergesagten Wert und dem beobachteten Wert des Outcome auf. Solange aber keine Abweichung vorliegt, findet keine Veränderung der Assoziationsgewichte statt. Das zweite Ereignis wird also nicht assoziiert. Das Ausbleiben eines Blocking-Effektes bei Annahme eines Gemeinsame-Ursache Modells können diese Theorien nicht modellieren, da Cues und Outcome in beiden Experimentalbedingungen identisch waren.

Probanden klar war. War eine Verwechslung zwischen Ursachen und Effekte möglich, dann näherten sich die Urteile bei den verschiedenen Kausalmodellen an (Waldmann, 2000).

Beim klassischen Blocking- und beim Overshadowing-Paradigma werden deterministische Zusammenhänge zwischen den Ereignissen verwendet. Die Frage ist, ob sich ein Einfluss von Kausalmodellen auch bei probabilistischen Zusammenhängen nachweisen lässt. So wurde Probanden ein Gemeinsamer-Effekt Modell vorgelegt mit der Aufgabe, die Stärke der beiden beteiligten Kausalzusammenhänge zu bestimmen. Die beiden Ursachen (bunte Flüssigkeiten) traten dabei sowohl individuell als auch zusammen auf und riefen den Effekt (Blühen von Pflanzen) nur mit einer bestimmten Wahrscheinlichkeit hervor (Spellman, 1996a,b). Im Gegensatz zum klassischen Overshadowing Paradigma lassen sich unter diesen Bedingungen die bedingten Kontingenzen der einzelnen Ursachen auf den Effekt bestimmen. Die Probanden sollten daher die normativ richtigen, auf die Abwesenheit der anderen Ursache bedingten Kontingenzen als Maß für den kausalen Einfluss heranziehen. Die Ergebnisse belegen, dass dies der Fall ist (Spellmann, 1996a,b). Gleiche Befunde wurden auch von anderen Forschern berichtet (Baker et al., 1993; Price & Yates, 1993).

In den oben genannten Experimenten gingen die Probanden von einem bestimmten Kausalmodell aus. Damit war die kausale Rolle aller beteiligten Ereignisse klar. Sie waren entweder Ursachen oder Effekte. Daneben gibt es aber auch Situationen, in denen das Kausalmodell nicht von vornherein bestimmt ist, beispielsweise wenn bei der Ermittlung eines Kausalzusammenhangs ein drittes Ereignis vorliegt, dessen kausaler Status noch nicht abschließend geklärt ist. In diesem Fall sollte das zusätzliche Ereignis nur berücksichtigt werden, wenn es für den Zieleffekt kausal relevant ist. Ansonsten kann es gemäß der Regel für die Bestimmung der Stärke eines Kausalzusammenhangs innerhalb eines Kausalmodells vernachlässigt werden (vgl. Kapitel 1.3). Die Frage ist daher, ob Probanden Ereignisse, die mit dem Zieleffekt kovariieren, aber nicht kausal relevant für diesen sind, konstant halten. Normativ gesehen, ist das Bestehen eines statistischen Zusammenhangs mit dem Zieleffekt nicht hinreichend dafür, das betreffende Ereignis konstant zu halten. In einem Experiment (Waldmann & Hagmayer, 1995, Experiment 1) wurden daher Probanden gebeten, die Wirkung einer Ursache (Bestrahlung von Früchten) auf einen Effekt (Qualität der Früchte) einzuschätzen. Zusätzlich lag stets ein weiteres Ereignis vor, das mit dem Effekt hoch korreliert war. Dieses Zusatzereignis war entweder kausal relevant (zwei unterschiedliche Sorten von Früchten) oder nicht relevant (per Zufall zugeordnete Person, welche die Proben untersucht hat). Die Urteile der Probanden zeigten, dass nur das kausal relevante

Zusatzereignis kontrolliert wurde. In diesem Fall wurden die normativ richtigen bedingten Kontingenzen als Indikator für Kausalität verwendet.

In verschiedenen Experimenten wurden Faktoren entdeckt, welche die Leistung begrenzten. Wichtig waren v.a. die Anforderungen an die Verarbeitungskapazität. So wurden zusätzliche, kausal relevante Faktoren nur dann berücksichtigt, wenn diese salient waren und leicht verarbeitet werden konnten (Fiedler et al., 2000; Schaller & O'Brian, 1992; Waldmann & Hagmayer, *subm.*). Ebenso führte eine Zusatzaufgabe (Zählen von Tönen während des Lernens) dazu, dass der Anteil der Probanden, welche das Kausalmodell beim Lernen richtig nutzten, zurückging (Waldmann, 2000). Eine Steigerung der Komplexität der gezeigten Ursache-Wirkungs-Zusammenhänge (gleichzeitiges Lernen mehrerer Kausalmodelle) hatte den gleichen Effekt (Waldmann, 2000).

Fassen wir die Befunde für das Lernen von Kausalmodellen zusammen. Die Ergebnisse belegen, dass Probanden sensitiv für Kausalmodelle sind. Probanden sind in der Lage, diese normativ richtig zu nutzen und den für das jeweilige Modell richtigen Indikator für Kausalität auszuwählen. Insbesondere wird beim Gemeinsamer-Effekt Modell der Einfluss der alternativen, konfundierenden Ursache kontrolliert. Die Probanden verwenden bei diesem Modell richtigerweise die bedingte Kontingenz als Indikator für die kausale Wirksamkeit. Daneben zeigt sich aber auch erneut ein Einfluss der Anforderungen an die Verarbeitungskapazität. Wie beim Erlernen von einfachen Kausalzusammenhängen ging der Anteil der Probanden, die normativ korrekt voringen, mit steigenden Anforderungen zurück.

Offene Fragen

Trotz dieser vielfältigen Belege für eine Berücksichtigung von Kausalmodellen bleiben noch Fragen offen. Nur in wenigen Studien wurden verschiedene Kausalmodelle direkt miteinander verglichen. Dabei haben sich die Forscher auf das Gemeinsame-Ursache und das Gemeinsamer-Effekt Modell beschränkt. Die Kettenmodelle mit ihren spezifischen Vorhersagen für die Bestimmung des ersten und des zweiten Kausalzusammenhangs wurden bisher vernachlässigt. Außerdem wurden in diesen vergleichenden Untersuchungen deterministische Zusammenhänge zwischen den Ereignissen verwendet. Es gilt daher die Befunde auf probabilistische Zusammenhänge zu erweitern. Hierzu wurde im Rahmen dieser Arbeit ein weiteres Experiment durchgeführt. Die Aufgabe der Probanden war es, die Stärke eines Kausalzusammenhangs zu bestimmen, der in verschiedene Kausalmodelle eingebettet war. Das Gemeinsame-Ursache Modell, das Gemeinsamer-Effekt Modell und zwei Kettenmodelle, wurden miteinander verglichen. Zusätzlich wurde als zweiter Faktor die Abstraktheit des verwendeten Materials manipuliert, um sicherzustellen, dass ein etwaiger

Einfluss der Kausalmodelle nicht auf das Vorwissen der Probanden über den gezeigten Sachverhalt zurückzuführen ist. Die unten stehende Abbildung zeigt das verwendete Versuchsdesign und die normativ richtigen Vorhersagen.

	Kettenmodell 1 $C \rightarrow A \rightarrow B$	Gemeinsame- Ursache Modell $C \leftarrow A \rightarrow B$	Gemeinsamer- Effekt Modell $A \rightarrow B \leftarrow C$	Kettenmodell 2 $A \rightarrow B \rightarrow C$
Konkreter Inhalt	<i>Unbedingte Kontingenz</i> ²² ΔP_{AB}	<i>Unbedingte Kontingenz</i> ΔP_{AB}	<i>Bedingte Kontingenz</i> ΔP_{AB-C}	<i>Unbedingte Kontingenz</i> ΔP_{AB}
Abstrakter Inhalt	<i>Unbedingte Kontingenz</i> ΔP_{AB}	<i>Unbedingte Kontingenz</i> ΔP_{AB}	<i>Bedingte Kontingenz</i> ΔP_{AB-C}	<i>Unbedingte Kontingenz</i> ΔP_{AB}

Abb. 2.6: Experiment zum Einfluss von Kausalmodellen auf die Bestimmung der Stärke eines Kausalzusammenhangs. Die Aufgabe ist, den Zusammenhang von A und B einzuschätzen. Die Einträge in den Zellen geben die normativ richtigen Indikatoren an.

Wenn die Probanden sensitiv für die verschiedenen Kausalmodelle sind, dann sollten bei konkreten wie abstrakten Inhalten die normativ richtigen Kontingenzen als Indikator für die Stärke des Kausalzusammenhangs der Ereignisse A und B verwenden, d.h., bei den beiden Kettenmodellen und dem Gemeinsame-Ursache Modell die unbedingte Kontingenz und beim Gemeinsamer-Effekt Modell die bedingte Kontingenz. Das Experiment und die Befunde werden in Abschnitt 3.1 ausführlich besprochen.

²² Dies gilt bei Annahme einer Markovkette (vgl. Kapitel 1.2)

2.2 Prognosen und Diagnosen

Die Aufgabe beim Erlernen von Kausalzusammenhängen ist es, aus einer Vielzahl von Lernerfahrungen die Stärke der Kausalrelation zu erschließen. Bei Prognosen und Diagnosen handelt es sich dagegen um Vorhersagen im Einzelfall. Dabei gilt es aus einer bestimmten Ursachekonstellation einen Effekt bzw. seine Wahrscheinlichkeit zu prognostizieren oder aus einer bestimmten Kombination von Effekten das Vorliegen bestimmter Ursachen zu diagnostizieren. Eine Voraussetzung für Prognosen und Diagnosen ist das Wissen über Kausalmodelle. Nur wenn ein solches Wissen gegeben ist, kann die Wahrscheinlichkeit eines Effektes bzw. einer Ursache unter gewissen Bedingungen zuverlässig geschätzt werden. Kausalmodelle erlauben es aber auch auf Ereignisse zu schließen, die mit dem Prädiktor nicht unmittelbar kausal verbunden sind. So können wir aus dem Stand des Barometers auf Regen schließen, obwohl weder der Stand des Barometers den Regen, noch der Regen das Barometer beeinflusst. Da aber beide von einer gemeinsamen Ursache, dem Luftdruck, abhängen, ist dieser Schluss erfolgreich. Die nun folgenden Teilkapitel widmen sich Prognosen (2.2.1), Diagnosen (2.2.2) und der Vorhersage nicht direkt kausal verbundener Ereignisse (2.2.3).

2.2.1 Prognose von Effekten

Bei Vorhersagen wird häufig auf Kausalmodelle zurückgegriffen. Wenn alle Wirkungen einer Ursache prognostiziert werden sollen, dann liegt dieser Aufgabe ein Gemeinsame-Ursache Modell zu Grunde. Durch das Auftreten der Ursache verändert sich die Wahrscheinlichkeit aller Effekte von $P(e)$ zu $P(e|u)$. Beispielsweise verringert die Einnahme eines Medikamentes die Wahrscheinlichkeit der damit behandelten Störung, erhöht aber die Wahrscheinlichkeit des Auftretens unerwünschter Nebenwirkungen. Diese bedingten Wahrscheinlichkeiten sind Bestandteile eines Gemeinsame-Ursache Modells und können daher direkt aus diesem abgelesen werden. Interessanter ist daher die Prognose eines Effektes aufgrund mehrerer verschiedener Ursachen, die in bestimmten Ausprägungen vorliegen. Da im Allgemeinen kein Ereignis nur von einer Ursache beeinflusst wird, sondern von einer Vielzahl von Faktoren, müssen wir häufig solche Vorhersagen machen. Diesen Aufgaben liegt ein Gemeinsamer-Effekt Modell zugrunde. Auf dessen Grundlage kann die Wahrscheinlichkeit des Effektes bei einer bestimmten Ursachekonstellation vorhergesagt werden. Dabei gibt es zwei Optionen. Die erste Möglichkeit ist, dass das Modell bereits Informationen über die Wahrscheinlichkeit des Effektes bei den jeweiligen Ursachekonstellationen beinhaltet. Damit kann die Wahrscheinlichkeit des Effektes je nach

Gegebenheiten direkt abgelesen werden. Diese Option verlangt aber sehr viel Wissen über mögliche Ursachekonstellationen und ihre Wirkungen. Eine sparsame andere Möglichkeit ist es daher, die Wahrscheinlichkeit des Effektes aus der Stärke der Kausalbeziehungen und den Basisraten der einzelnen Ursachen zu erschließen. Sind die Ursachen unabhängig und wirken auch unabhängig voneinander auf den Effekt, dann kann das ‚noisy-or Schema‘ (Pearl, 1988; vgl. Kapitel 1.2) angewandt werden. Nach diesem gilt:

$$P(E|U1,U2) = (1 - (1-P(E|U1)) \cdot (1-P(E|U2))). \quad (\text{noisy-or})$$

Liegen beide Ursachen in einem konkreten Fall vor, ergibt sich die gesuchte Wahrscheinlichkeit als $P(e|u1.u2) = 1 - (1-P(e|u1)) \cdot (1-P(e|u2))$. Auf diesem Weg können wir beispielsweise bestimmen, mit welcher Wahrscheinlichkeit ein rauchender Bauarbeiter, der in der Vergangenheit ungeschützt mit Asbest gearbeitet hat, Lungenkrebs bekommt. Nehmen wir an, dass die Wahrscheinlichkeit durch Rauchen Lungenkrebs zu bekommen bei $P(\text{Lungenkrebs}[lk] | \text{Rauchen}[r]) = 0,20$ liegt und bei Asbestbelastung $P(\text{Lungenkrebs}[lk] | \text{Asbest}[a]) = 0,40$ beträgt, dann ergibt sich als Wahrscheinlichkeit für Lungenkrebs für den Arbeiter, $P(lk | r.a) = 1 - (1-P(lk|r)) \cdot (1-P(lk|a)) = 1 - 0,8 \cdot 0,6 = 0,52$. Die Wahrscheinlichkeit steigt also durch die Doppelbelastung an. Sofern aber keine Interaktion der beiden Ursachen gegeben ist, ist die resultierende Wahrscheinlichkeit stets kleiner als die Summe der Einzelwahrscheinlichkeiten.

Empirische Befunde

Ob Personen in der Lage sind, für bestimmte Ursachenkombinationen Prognosen abzuleiten, wurde unter anderem in Experimenten zur Urteilsrevision untersucht (Plach, 1998). In diesen wurden Probanden zunächst gebeten, alle bedingten Wahrscheinlichkeiten und Basisraten eines Kausalmodells zur Entstehung von Verkehrsstaus anzugeben. Danach sollten sie für verschiedene Kombinationen von Ursachenfaktoren schätzen, mit welcher Wahrscheinlichkeit auf der vor ihnen liegenden Strecke ein Stau zu erwarten ist. Zu diesen Ursachenfaktoren gehörten u.a. die Tageszeit (Rushhour vs. nachmittags) und die Bevölkerungsdichte (ländlicher Raum vs. Ballungsraum). Das Ergebnis war, dass sich die angegebene Wahrscheinlichkeit für einen Stau individuell durch die geschätzten bedingten Wahrscheinlichkeiten mit Hilfe eines Kausalmodells vorhersagen ließ. Zudem konnte gezeigt werden, dass ein Bayesianisches Modell die Schätzungen der Probanden besser vorhersagte als ein lineares Regressionsmodell.

2.2.2 Diagnose von Ursachen

Die einfachste Form der Diagnose ist es, aus dem Vorliegen eines Effektes auf das Vorliegen einer Ursache zu schließen. Die Wahrscheinlichkeit, dass die Ursache vorliegt, kann dabei mit Hilfe der Bayesformel bestimmt werden. Nach dieser ist die Wahrscheinlichkeit der Ursache, wenn der Effekt vorliegt (die a posteriori Wahrscheinlichkeit) gleich dem Produkt aus der Wahrscheinlichkeit des Effektes bei Vorliegen der Ursache und der Basisrate der Ursache (a priori Wahrscheinlichkeit) geteilt durch die absolute Wahrscheinlichkeit des Effektes.

$$P(\text{Ursache} \mid \text{Effekt}) = P(\text{Effekt} \mid \text{Ursache}) \cdot P(\text{Ursache}) / P(\text{Effekt}) \quad (\text{Bayestheorem})$$

Die absolute Wahrscheinlichkeit des Effektes ergibt sich dabei als

$$P(\text{Effekt}) = P(\text{Effekt} \mid \text{Ursache}) \cdot P(\text{Ursache}) + P(\text{Effekt} \mid \sim \text{Ursache}) \cdot P(\sim \text{Ursache})$$

Die Wahrscheinlichkeit der Ursache lässt sich also über die bedingten Wahrscheinlichkeiten der Kausalzusammenhänge und die Wahrscheinlichkeit der Ursache bestimmen. Zum Beispiel kann auf diesem Wege ermittelt werden, wie groß die Wahrscheinlichkeit für eine Frau ist, Brustkrebs zu haben, wenn eine Mammographie einen positiven Befund liefert (vgl. Jungermann, 1998, S. 180-181). Die a priori Wahrscheinlichkeit für Brustkrebs ist bei Frauen sehr gering ($P(u) = 0,01$), die Sensitivität ($P(e|u) = 0,80$) und die Spezifität ($P(\sim e|\sim u) = 0,90$) des Testverfahrens sind aber sehr hoch. Setzt man diese Zahlen in die Bayesformel ein, dann ist unter diesen Bedingungen die a posteriori Wahrscheinlichkeit, dass bei einem positiven Befund tatsächlich Brustkrebs gegeben ist, immer noch relativ gering ($P(u|e) = 0,075$). Der Grund hierfür ist die geringe Basisrate der Ursache.

Das Bayestheorem lässt sich auch auf komplexere Kausalmodelle anwenden. Wird ein Effekt von mehreren verschiedenen, unabhängigen Ursachen hervorgerufen (Gemeinsamer-Effekt Modell), dann kann die Wahrscheinlichkeit jeder Ursache bei Vorliegen des Effektes über die Bayesformel bestimmt werden. Wenn es zum Beispiel um die Frage geht, ob für einen Verkehrsunfall ein Versagen der Bremsen oder Trunkenheit am Steuer die Ursache waren, dann kann die a posteriori Wahrscheinlichkeit für jede der beiden möglichen Ursachen separat berechnet werden. Liegt jedoch ein Gemeinsame-Ursache Modell vor, bei dem eine Ursache mehrere Effekte hervorruft, dann führt die einfache Bayesformel nicht weiter. Diese kann aber entsprechend erweitert werden. Es gilt dann, dass

$$P(U|E_1, E_2) = P(E_1, E_2|U) \cdot P(U) / P(E_1, E_2) \quad (\text{Bayestheorem für zwei Ursachen})$$

Da die Ursache unabhängig voneinander auf beide Effekte wirkt, vereinfacht sich die Berechnung auf

$$P(u|e_1, e_2) = \frac{P(e_1|u) \cdot P(e_2|u) \cdot P(u)}{[P(e_1|u) \cdot P(e_2|u) \cdot P(u) + P(e_1|\sim u) \cdot P(e_2|\sim u) \cdot P(\sim u)]}$$

Nehmen wir zum Beispiel die Diagnose von AIDS. Es gibt derzeit zwei Testverfahren, die zur Diagnose eingesetzt werden. Dies ist einerseits der ELISA Test (enzyme-linked immunosorbent assay) andererseits der Western Blot Test. Beide Tests weisen eine hohe Sensitivität ($P(e|u)$) und Spezifität ($P(\sim e|\sim u)$) auf. Die unten stehende Abbildung zeigt das zugehörige Kausalmodell und die bedingten Wahrscheinlichkeiten (nach Clemen, 1996, S.259-262).

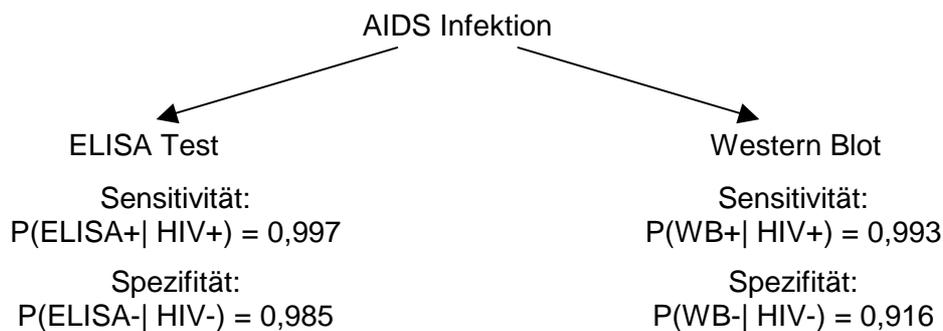


Abb. 2.7: Kausalmodell für die Diagnose von AIDS
(Werte nach Clemen, 1996, S.259-262)

Die Basisrate für AIDS betrug Ende der neunziger Jahre in den USA ca. $P(\text{AIDS}+) = 0,0038$. Mit diesen Daten kann nun auf Grundlage des Gemeinsame-Ursache Modells leicht die Wahrscheinlichkeit einer AIDS Infektion je nach Testergebnis ermittelt werden. Folgende Wahrscheinlichkeiten ergeben sich.

$$\begin{aligned}
 P(\text{HIV}+|\text{ELISA}+) &= P(\text{ELISA}+|\text{HIV}+) \cdot P(\text{HIV}+) / [P(\text{ELISA}+|\text{HIV}+) \cdot P(\text{HIV}+) \\
 &\quad + P(\text{ELISA}+|\text{HIV}-) \cdot P(\text{HIV}-)] = 0,20 \\
 P(\text{HIV}+|\text{WB}+) &= P(\text{WB}+|\text{HIV}+) \cdot P(\text{HIV}+) / [P(\text{WB}+|\text{HIV}+) \cdot P(\text{HIV}+) \\
 &\quad + P(\text{WB}+|\text{HIV}-) \cdot P(\text{HIV}-)] = 0,043 \\
 P(\text{HIV}+|\text{ELISA}+.\text{WB}+) &= P(\text{ELISA}+.\text{WB}+|\text{HIV}+) / P(\text{ELISA}+.\text{WB}+) \\
 &= P(\text{ELISA}+|\text{HIV}+) \cdot P(\text{WB}+|\text{HIV}+) \cdot P(\text{HIV}+) / \\
 &\quad P(\text{ELISA}+.\text{WB}+) \\
 &= 0,75
 \end{aligned}$$

Was diese Berechnungen deutlich machen, ist, dass die Wahrscheinlichkeit wirklich mit AIDS infiziert zu sein, bei nur einem positiven Testergebnis sehr gering ist. Sind aber beide Tests positiv, dann steigt die Wahrscheinlichkeit auf 75%

Empirische Befunde

Diagnosen bei einzelnen Kausalzusammenhängen wurden in einer Vielzahl von Experimenten untersucht. Dazu wurden Probanden das Brustkrebsproblem und ähnliche Aufgabenstellungen vorlegt. Die Frage in diesen Untersuchungen war, ob Probanden zu Bayesianischem Denken fähig sind (u.a. Kahnemann, Slovic & Tversky, 1982). Als Daten

erhielten die Probanden Informationen über die Basisrate der Ursache und die bedingten Wahrscheinlichkeiten in Bezug auf den Kausalzusammenhang. Aufgabe war es, die Wahrscheinlichkeit der Ursache einschätzen, wenn der Effekt gegeben ist. Die Ergebnisse waren ernüchternd. Die tatsächliche Wahrscheinlichkeit der Ursache bei Vorliegen des Effektes wurde erheblich überschätzt (Kahnemann, Slovic & Tversky, 1982) und zwar sowohl von Laien wie von Experten (Eddy, 1982). Dieser Befund wurde dadurch erklärt, dass die Probanden die niedrige Basisrate nicht berücksichtigen. Daher ist dieses Phänomen auch als „base rate neglect“ in die Literatur eingegangen (Dawes, 1988; Jungermann et al., 1998).

In anderen Studien konnte aber gezeigt werden, dass Probanden durchaus in der Lage sind, Basisraten zu nutzen und normativ richtige Urteile zu fällen. Sie können dies, wenn die Daten nicht in Form von Wahrscheinlichkeiten, sondern als absolute Häufigkeiten präsentiert werden (Gigerenzer & Hoffrage, 1995). Beim Mammographieproblem wurden die Probanden nun informiert, dass von 1000 Frauen, 10 an Brustkrebs leiden. Eine Mammographie liefert bei 8 dieser 10 Frauen einen positiven Befund. Von den restlichen 990 gesunden Frauen haben 99 einen positiven Befund. Wurden die Daten in dieser Form dargeboten, dann kam die Mehrheit der Probanden zu einer korrekten Antwort.

Andere Studien haben sich mit Diagnosen bei komplexen Kausalmodellen beschäftigt. Grundlage der Diagnosen war dabei ein komplexes Gemeinsamer-Effekt Modell mit mehreren verschiedenen möglichen Ursachen. Die Aufgabe der Probanden war es anzugeben, welche der möglichen, in Frage kommenden Ursachen der Auslöser für das Effektereignis war. Studien dieser Art wurden in verschiedenen psychologischen Teildisziplinen durchgeführt.

In der Sozialpsychologie beschäftigt man sich seit langem mit der Kausalattribution von Verhalten. Dabei geht es um die Frage, wie Probanden bei der Ermittlung der Ursache von Verhaltensweisen (z.B. „Tom schlägt Dick während der großen Pause“) vorgehen. Als mögliche Ursachen wurden die Situation (große Pause), die Reizgegebenheit (Dick) und die Person (Tom) vorgegeben. Für ihre Diagnose erhielten die Versuchsteilnehmer Informationen darüber, ob sich auch andere Personen gegenüber dem Stimulus gleich verhalten (Konsensus), ob sich der Protagonist gegenüber dem Stimulus in anderen Situationen gleich verhält (Konsistenz) und ob sich der Protagonist gegenüber anderen Stimuli gleich verhält (Distinktheit). Diese Informationen wurden in Form von groben, nicht numerischen Aussagen über die Kovariation der entsprechenden Ursache mit dem Effekt vorgegeben (z.B. „Tom schlägt auch andere Kinder“). Die Abbildung auf der folgenden Seite stellt die gegebenen Informationen formal dar. So besagt beispielsweise ein hoher Konsensus, dass die Wahrscheinlichkeit des Verhaltens bei gleichem Stimulus in der

gleichen Situation beim Protagonisten und einer anderen Person gleich ist. Damit ist das Verhalten in dieser Situation bei diesem Stimulus unabhängig vom Protagonisten.

Konsensus		Konsistenz		Distinktheit	
hoch	niedrig	hoch	niedrig	hoch	niedrig
$P(V St.P.Si)=P(V St.\sim P.Si)$	$P(V St.P.Si)\neq P(V St.\sim P.Si)$	$P(V St.P.Si)=P(V St.P.\sim Si)$	$P(V St.P.Si)\neq P(V St.P.\sim Si)$	$P(V St.P.Si)\neq P(V \sim St.P.Si)$	$P(V St.P.Si)=P(V \sim St.P.Si)$

Abb. 2.8: Informationen in Experimenten zur Kausalattribution
V = zu erklärendes Verhalten, St = Stimulus, P = Protagonist, Si = Situation

Es konnte gezeigt werden, dass Probanden bei der Attribution die Kovariation der einzelnen Ursachen mit dem Effekt berücksichtigen und dabei alle drei Informationsquellen heranziehen (Kelley, 1967, 1973). So wird beispielsweise auf den Stimulus als Ursache attribuiert, wenn sowohl der Konsens als auch die Konsistenz und die Distinktheit hoch sind. In unserem Beispiel bedeutet dies, dass Tom nur Dick schlägt und auch andere Kinder Dick schlagen.

Dennoch ist Kritik an diesem Ansatz der Kausalattribution angebracht (vgl. auch Cheng & Novick, 1990, für eine ausführliche Kritik). Um wirklich herauszufinden, wodurch der Effekt hervorgerufen wurde, werden weitere Informationen benötigt. Wir hatten oben gesehen, dass bei einem Gemeinsamer-Effekt Modell die bedingten Kontingenzen indikativ für die Stärke des Kausalzusammenhangs sind. Für die Bestimmung dieser bedingten Kontingenzen sind aber die bedingten Wahrscheinlichkeiten des Effektes bei allen denkbaren Ursachenkonstellationen notwendig. Zum Beispiel wird die Wahrscheinlichkeit benötigt, mit der andere Kinder auch außerhalb der konkreten Situation das Verhalten zeigen ($P(V|\sim St.\sim P.\sim Si)$). Es könnte ja sein, dass die Verhaltensweise andere zu schlagen insgesamt sehr häufig vorkommt und daher dieses Verhalten weder auf die Situation noch den Stimulus noch den Protagonisten zurückzuführen ist. Auf Basis dieser vollständigen Daten kann der Einfluss aller Ursachen und Kombinationen von Ursachen genau bestimmt werden. So ergeben sich z.B. für den Stimulus als Ursache vier verschiedene Kontingenzen

$$\Delta P_1 = P(V|St.P.Si) - P(V|\sim St.P.Si), \Delta P_2 = P(V|St.P.\sim Si) - P(V|\sim St.P.\sim Si),$$

$$\Delta P_3 = P(V|St.\sim P.Si) - P(V|\sim St.\sim P.Si), \Delta P_4 = P(V|St.\sim P.\sim Si) - P(V|\sim St.\sim P.\sim Si).$$

Nur wenn all diese Kontingenzen positiv und gleich sind, ist Dick als Stimulus die alleinige Ursache dafür gewesen, dass Tom ihn geschlagen hat. Durch die Bestimmung der bedingten Kontingenzen für alle Ursachenkonstellationen werden auch eventuelle Interaktionen sichtbar. Wenn die vier bedingten Kontingenzen voneinander abweichen, dann liegt eine Interaktion von verschiedenen Ursachenfaktoren vor (vgl. Cheng & Novick, 1990). Experimente haben gezeigt, dass Personen diese weitergehenden Informationen nutzen und

zu anderen Urteilen gelangen (Cheng & Novick, 1990). So attribuieren Versuchspersonen durchaus auch auf mehrere Faktoren gleichzeitig. Probanden sind also in der Lage, noch normativer vorzugehen, als dies in Kelleys Modell vorgesehen ist.

Ein weiteres Problem ist, dass es sich bei der Kausalattribution von Verhalten zwar um Diagnosen handelt (die Ursache für ein Verhalten soll ermittelt werden), die gegebenen Daten aber nicht hinreichend für ein normatives Vorgehen sind. So ist eine Anwendung des Bayestheorems aufgrund fehlender Basisraten nicht möglich.²³ Was die Studien aber zeigen, ist, dass in diesem Fall der Faktor, welcher die stärkste Wirkung hat bzw. die höchste bedingte Kontingenz aufweist, als Ursache des Effektes diagnostiziert wird.

Auch Untersuchungen zum kausalen Denken haben sich mit der Frage beschäftigt, welche Informationen bei Diagnosen genutzt werden. Dazu wurden Probanden alltägliche Ereignisse (u.a. ein Autounfall und eine allergische Reaktion) vorgelegt und sie wurden gebeten die Ursache zu diagnostizieren. Dazu durften die Probanden beliebige Fragen an den Versuchsleiter stellen. Die Fragen wurden nach einem Klassifikationssystem ausgewertet. Dabei stellte sich heraus, dass nur sehr selten Fragen nach der Kovariation möglicher Ursachen mit dem Effekt gestellt wurden (<10%). Die überwiegende Mehrheit (ca. 2/3) der Fragen bezog sich darauf, ob mögliche Ursachen in dem spezifischen Fall vorlagen oder nicht (Ahn et al., 1995). Die Probanden hatten offensichtlich spezifische Hypothesen und überprüften diese. Die dabei nachgefragten Ursachen waren nach allgemeinem Vorwissen solche, die mit hoher Wahrscheinlichkeit den Effekt hervorrufen und häufig auftreten. Zum Beispiel wurde bei der Ermittlung der Ursache des Autounfalls häufig nach Trunkenheit des Fahrers gefragt, aber selten nach einem Hirnschlag, obwohl auch dieser mit hoher Wahrscheinlichkeit zu einem Unfall geführt hätte. Dieser Befund deutet darauf hin, dass bei Diagnosen bezüglich bekannter Sachverhalte, Personen auf ihr Kausalmodell des Sachverhaltes zurückgreifen und ermitteln, welche der Ursachen innerhalb des Modells vorliegen. Das Wissen über die Wirksamkeit der jeweiligen Ursache kann dabei als vorhanden vorausgesetzt werden.²⁴

²³ Möglich wäre es, die Likelihood Ratio zu bestimmen. Diese ist definiert als der Quotient aus der bedingten Wahrscheinlichkeit des Effektes bei Vorliegen der Ursache und der bedingten Wahrscheinlichkeit des Effektes ohne Vorliegen der Ursache. Dieser Quotient ist proportional dem Quotienten aus der a posteriori Wahrscheinlichkeit der Ursache, nachdem der Effekt eingetreten ist, und der a posteriori Wahrscheinlichkeit, dass die Ursache nicht vorlag, nachdem der Effekt eingetreten ist. Die Likelihood Ratio gibt an, um wie viel es wahrscheinlicher wird, dass die Ursache vorlag, wenn der Effekt aufgetreten ist. Eine absolute Aussage ermöglicht sie aber nicht.

²⁴ Vertreter eines mechanistischen Ansatzes von Kausalität werten diese Ergebnisse allerdings als Beweis dafür, das Wissen um Mechanismen wichtiger ist als das Wissen über Kovariation (Ahn et al., 1995). Meines Erachtens ist diese Schlussfolgerung nicht gerechtfertigt (vgl. auch Cheng, 1997, für einen integrativen Ansatz)

Eine Reihe von Experimenten hat sich mit Diagnosen bei komplexen Kausalmodellen beschäftigt und untersucht, ob bei diesen ein base-rate neglect nachgewiesen werden kann. Hierzu wurde ein kausales Strukturmodell verwendet, bei dem zwei Ursachen jede für sich einen Effekt und beide zusammen einen dritten Effekt deterministisch hervorriefen, d.h., eine Ursache U_1 führte zu den Effekten E_1 und E_2 und eine Ursache U_2 zu den Effekten E_2 und E_3 (Reips, 1997). Die Ereignisse bildeten also eine M-förmige Struktur. Die Basisraten der beiden Ursachen waren stark unterschiedlich. Abbildung 2.9 zeigt den Grundaufbau des Experimentes.

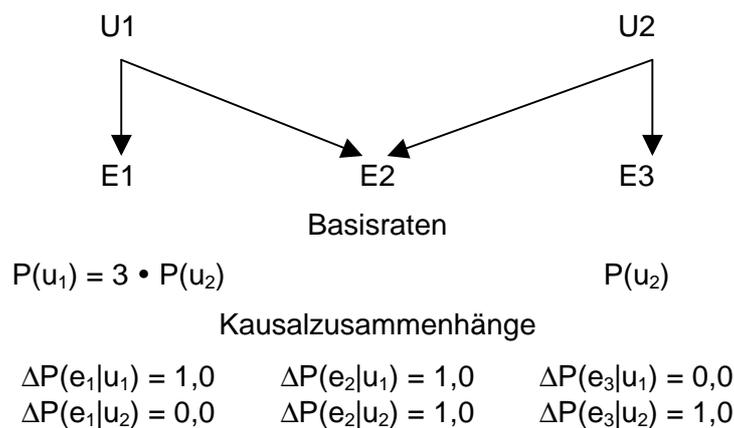


Abb. 2.9: Aufbau eines Experiment zur Diagnose bei komplexen Kausalmodellen (vgl. Reips, 1997)

Nachdem sie das Kausalmodell erlernt hatten, sollten die Probanden die Wahrscheinlichkeit der beiden Ursachen einschätzen, wenn nur einer der Effekte vorlag. Sie sollten also $P(u_x|e_y)$ angeben. Interessant ist dabei die Schätzung der bedingten Wahrscheinlichkeit der beiden Ursachen, wenn allein der zweite Effekt vorliegt. Wenn die Probanden die Basisraten berücksichtigen, was sie nach dem Bayestheorem und dem Kausalmodell tun sollten, dann müssten sie die Wahrscheinlichkeit für die erste Ursache deutlich höher bewerten. Das gefundene Ergebnis bestätigte diese Vermutung. Weitere Experimente haben gezeigt, dass die Nutzung der Basisrate dabei unabhängig von der Lernrichtung ist. Die Basisrate wurde sowohl nach prädiktivem Lernen (von der Ursache zum Effekt) als auch nach diagnostischem Lernen (vom Effekt zur Ursache) berücksichtigt. Bei komplexeren Aufgabenstellungen mit mehreren Kausalmodellen, die gleichzeitig zu erlernen waren, ging die Basisratennutzung zurück. Sie zeigte sich nur noch nach diagnostischem Lernen. Bei dieser Form des Lernens bekommen die Probanden direkt Informationen über die a posteriori Wahrscheinlichkeit, welche die unterschiedlichen Basisraten der Ursachen direkt beinhaltet. Im Gegensatz dazu werden beim prädiktiven Lernen die bedingten Wahrscheinlichkeiten die Kausalzusammenhänge betreffend erlernt. Bei späteren Diagnosen

muss dieses Wissen dann mit dem Wissen über die Basisraten richtig kombiniert werden. Dies scheint bei hoher Komplexität nicht mehr möglich zu sein (Reips, 1997). Eine Sensitivität für die Basisrate konnte nach diagnostischem Lernen auch noch bei Strukturen mit 15 Ereignissen nachgewiesen werden (Reips, 1997). Probanden haben also die Kompetenz, auch bei sehr komplexen Kausalmodellen normativ richtige Diagnosen zu stellen.

Die bisher dargestellten Studien hatten sich auf Kausalmodelle mit zwei Ebenen, eine Ursache- und eine Effektebene beschränkt. Fast alle Sachverhalte in der Welt bilden aber komplexe kausale Netzwerke, in denen Ursachen intermittierende Ereignisse hervorrufen, und diese wiederum bestimmte Effekte. In Experimenten zur Urteilsrevision wurden Diagnosen bei solchen komplexen Kausalmodellen untersucht (Plach, 1998). Dabei wurden Probanden gebeten, auf der Grundlage eines komplexen Modells zur Entstehung von Verkehrsstaus die Wahrscheinlichkeit eines Staus bei verschiedenen Gegebenheiten zu diagnostizieren. Als Evidenzen erhielten sie Informationen über Effekte von Staus, z.B. entsprechende Radiomeldungen. Es zeigte sich, dass sich die Urteile der Probanden über die Bayesformel für komplexe Kausalmodelle vorhersagen ließen.

2.2.3 Vorhersagen nicht direkt kausal verbundener Ereignisse

Bei Prognosen wie Diagnosen geht es um die Vorhersage von Ereignissen, welche direkt kausal mit dem Prädiktoreignis verbunden sind. Kausalmodelle erlauben aufgrund ihrer strukturellen Implikationen aber auch die Vorhersage von Ereignissen, die mit dem Prädiktoreignis nicht direkt kausal verbunden sind. So impliziert das Gemeinsame-Ursache Modell einen Zusammenhang zwischen den beiden Effekten der einen Ursache. Daher kann aus dem Vorliegen des einen Effektes auf das Vorliegen eines anderen Effektes geschlossen werden. Diesen Sachverhalt machen wir uns beispielsweise in der Medizin zu Nutze. So schließen Ärzte aus dem Vorliegen äußerer Krankheitssymptome nicht nur auf das Vorliegen der Krankheit, sondern auch auf weitere, später eintretende Krankheitsfolgen. Zeigt ein Patient die Symptome einer Grippeerkrankung, dann schließen wir daraus, dass ein Kontakt mit ihm die Wahrscheinlichkeit einer Infektion erhöht und vermeiden daher einen direkten Umgang. Das Gemeinsamer-Effekt Modell impliziert keinen Zusammenhang zwischen den verschiedenen Ursachen desselben Effektes. Jedoch impliziert es eine auf das Vorliegen des Effektes bedingte Abhängigkeit. Liegt der Effekt vor und es ist bereits eine der Ursachen gegeben, dann sinkt die Wahrscheinlichkeit der anderen Ursache. Auch diese strukturelle Implikation machen wir uns in der Medizin zu Nutze. Wenn ein Patient eine

schwere Lungenentzündung hat und wir herausgefunden haben, dass er an Tuberkulose leidet, dann werden wir ihn nicht auch noch gegen andere mögliche virale oder bakteriologische Infektionen behandeln.

Die strukturellen Implikationen von Kausalmodellen erlauben auch die Vorhersage noch weiter voneinander entferntere Ereignisse innerhalb eines Kausalmodells. Folgendes Beispiel zeigt, wie zwei Ereignisse, die weder eine gemeinsame Ursache noch einen gemeinsamen Effekt miteinander haben, dennoch unter bestimmten Umständen zur gegenseitigen Vorhersage genutzt werden können (aus Jensen, 1997). Wir wollen dabei von folgendem Fall ausgehen: Die Herren Holmes und Watson wohnen direkt nebeneinander. Wenn es regnet, dann wird der Rasen sowohl im Garten von Holmes wie von Watson nass. Holmes besitzt aber zusätzlich einen Rasensprenger zur Bewässerung seines Rasens. Die Abbildung 2.10 zeigt das zugehörige Kausalmodell.

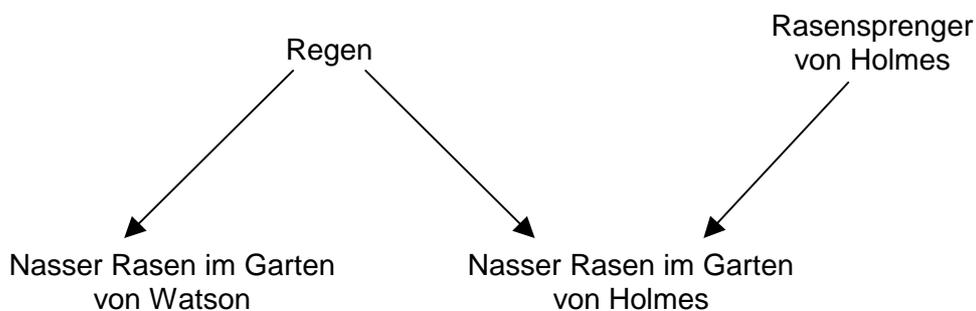


Abb. 2.10: Beispiel für die Vorhersage nicht direkt kausal verbundener Ereignisse (Nach Jensen, 1997, S.22 ff.)

Eines Morgens stellt Holmes beim Aufstehen fest, dass sein Rasen nass ist. Hierfür gibt es zwei mögliche Erklärungen, entweder es hat geregnet oder er hat vergessen, seinen Rasensprenger abzuschalten. Nun blickt Holmes über den Gartenzaun und bemerkt, dass auch bei seinem Nachbarn Watson der Rasen nass ist. Daraus folgert er, dass es wohl geregnet haben muss. Was bedeutet dies aber für den Rasensprenger von Holmes? Hat er vergessen ihn abzuschalten oder nicht? Unsere Intuition sagt uns, dass er den Rasensprenger wohl nicht vergessen hat. Das bedeutet, wir können in dieser Situation aus dem Wissen über den nassen Rasen bei Watson, eine Diagnose für den Rasensprenger von Holmes machen, und dies, obwohl der Rasensprenger von Holmes keinerlei kausalen Einfluss auf den Rasen von Watson hat. Dass dies so ist, folgt direkt aus den strukturellen Implikationen des zugehörigen Kausalmodells. Die Beobachtung des nassen Rasens bei Holmes führt zunächst zu einem Anstieg der Wahrscheinlichkeit beider möglicher Ursachen innerhalb dieses Gemeinsamer-Effekt Modells. Der nasse Rasen bei Watson erhöht dann nochmals die Wahrscheinlichkeit, dass es geregnet hat. Steigt aber innerhalb eines Gemeinsamer-Effekt Modells die Wahrscheinlichkeit einer Ursache an, wobei der Effekt

sicher eingetreten ist, dann sinkt wie oben beschrieben die Wahrscheinlichkeit der anderen Ursache.²⁵ Das Beispiel zeigt, wie die strukturellen Implikationen des Gemeinsame-Ursache und des Gemeinsamer-Effekt Modells auch auf die Vorhersage von Ereignissen bei komplexeren Modellen angewandt werden können.

Empirische Befunde

Mir sind derzeit keine Studien bekannt, die sich direkt mit der Vorhersage nicht kausal verbundener Ereignisse befasst haben. Es gibt aber indirekte Hinweise dafür, dass Personen sensitiv für die strukturellen Implikationen der Modelle bei Diagnosen sind. So trat in Studien zur Kausalattribution das Phänomen der Abwertung alternativer Ursachen auf (Kelley, 1972). Weiß der Proband, dass der Effekt gegeben ist und eine der möglichen Ursachen vorliegt, dann wird die Wahrscheinlichkeit der anderen Ursachen als eher gering eingeschätzt. Dieses Antwortverhalten entspricht exakt den strukturellen Implikationen des Gemeinsamer-Effekt Modells bei mehreren unabhängigen Ursachen.²⁶

Fassen wir die Erkenntnisse zu Prognosen und Diagnosen zusammen. Die Ergebnisse einer Vielzahl von Studien aus verschiedenen psychologischen Teildisziplinen belegen, dass Personen grundsätzlich in der Lage sind, normativ richtige Prognosen und Diagnosen auf der Basis von Kausalmodellen zu stellen. Sogar bei komplexen Modellen konnte unter gewissen Bedingungen ein normativ richtiges Vorgehen nachgewiesen werden. Insbesondere bei vertrauten Sachverhalten (Vorhersage von Verkehrsstaus) erzielten Probanden eine gute Leistung.

Des Weiteren wurden erneut vereinfachende Heuristiken gefunden. Bei den Aufgaben zur Kausalattribution waren Personen trotz unvollständiger Information in der Lage sinnvolle Schlussfolgerungen zu ziehen. Bei Diagnosen bezüglich vertrauter Ereignisse vereinfachten die Personen ihr Vorgehen dadurch, dass sie nur das Vorliegen der wahrscheinlichsten Ursachen überprüften. Daneben spielten wiederum die Anforderungen an die Verarbeitungskapazität eine wichtige Rolle. Bei sehr komplexen Sachverhalten ging die Leistung insbesondere bei Diagnosen deutlich zurück.

Alles in allem zeigt sich also ein sehr ähnliches Bild wie beim Erlernen von Kausalzusammenhängen, obwohl es sich bei Prognosen und Diagnosen immer um die

²⁵ vgl. Jensen, 1997, S.22 ff für eine Berechnung dieses Beispiels

²⁶ Nach Kelley ist das Abwertungsprinzip mit dem kausalen Schema mehrerer hinreichender Ursachen verbunden (Kelley, 1972). Weshalb es bei diesem Schema aber zur Abwertung von alternativen Ursachen kommt, kann über den Schemabegriff nicht erklärt werden. Hierzu werden Kausalmodelle mit ihren spezifischen strukturellen Implikationen benötigt. Kausalmodelle können auch das Aufwertungsprinzip beim Schema mehrerer notwendiger Ursachen erklären. In diesem Fall haben wir es mit Ursachen zu tun, die nur in Interaktion miteinander zu einem Effekt führen. Dadurch steigt die Wahrscheinlichkeit des Effektes bei mehreren Ursachen nicht linear an.

Vorhersage von Einzelereignissen bzw. deren Wahrscheinlichkeit handelt und nicht um die Bestimmung von Zusammenhängen.

Offene Fragen

Bisher haben sich meines Wissens noch keine Untersuchungen direkt mit der Frage beschäftigt, ob Probanden in der Lage sind, auf der Grundlage von Kausalmodellen Ereignisse vorherzusagen, die nicht direkt kausal mit dem Prädiktor verbunden sind.²⁷ Daher wurde im Rahmen dieser Arbeit eine Reihe von Experimenten zu dieser Fragestellung durchgeführt.

In diesen erlernten die Probanden zunächst separat voneinander zwei Kausalzusammenhänge, welche entweder ein Gemeinsame-Ursache oder ein Gemeinsamer-Effekt Modell bildeten. Das separate Lernen stellte sicher, dass sie keinen Zusammenhang zwischen den beiden vermeintlichen Effekten (Gemeinsame-Ursache Modell) bzw. den beiden Ursachen (Gemeinsamer-Effekt Modell) beobachten konnten. Danach wurden die Probanden gebeten, aus dem Vorliegen des einen Ereignisses auf das Vorliegen des nicht direkt kausal mit diesem verbundenen Ereignisses zu schließen. Die unten stehende Abbildung zeigt das grundlegende Versuchsparadigma.

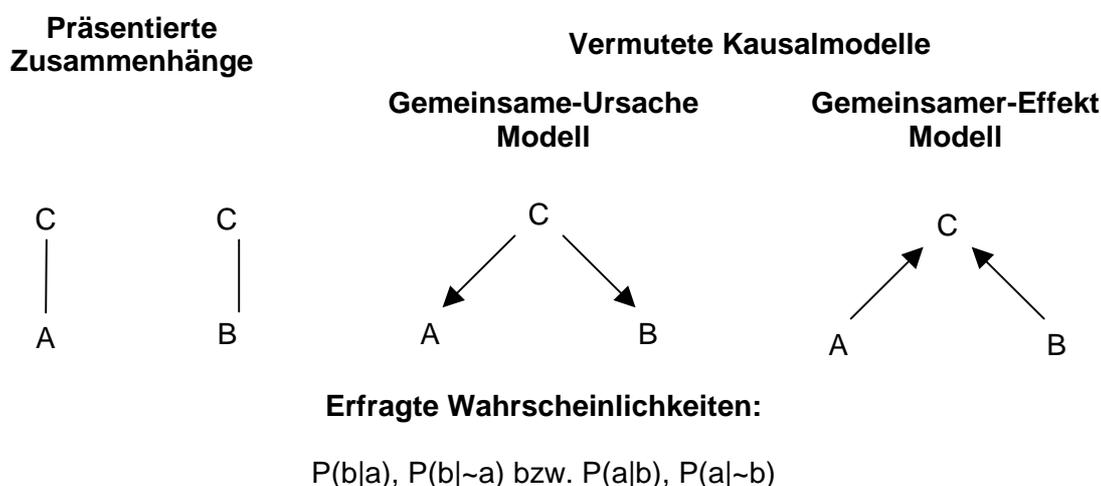


Abb. 2.11: Experiment zum Einfluss von Kausalmodellen auf die Vorhersage nicht direkt kausal verbundener Ereignisse

Wenn die Probanden sensitiv für die strukturellen Implikationen der beiden Kausalmodelle sind, dann sollten sie bei einem Gemeinsame-Ursache Modell aus dem Vorliegen des einen Effektes schließen, dass die Wahrscheinlichkeit des anderen Effektes

²⁷ Diese Aussage bezieht sich nur auf die Vorhersage von Ereignissen innerhalb von Kausalmodellen. Es liegen viele Untersuchungen zur Vorhersage von unbekanntem Merkmalen aus bekannten Merkmalen z.B. über Regressions- oder brunswikianische Linsenmodelle vor (vgl. Gigerenzer et al., 2000)

zunimmt ($P(b|a) > P(b) > P(b|\sim a)$). Bei einem Gemeinsamer-Effekt Modell sollte aber die Einschätzung der Wahrscheinlichkeit der zweiten Ursache trotz Vorliegen der ersten Ursache unverändert bleiben ($P(b|a) = P(b|\sim a) = P(b)$). Die ausführliche Beschreibung der Experimente und die Befunde sind in Kapitel 3.3 zu finden.

2.3 Prüfen von Kausalhypothesen

Es gibt zwei grundlegende Arten von Kausalhypothesen. Die einen beziehen sich auf kausale Zusammenhänge. Diese Hypothesen behaupten einen spezifischen Zusammenhang einer bestimmten Stärke und Richtung zwischen zwei Ereignissen. Ein Beispiel für eine solche Hypothese ist, dass Rauchen mit einer Wahrscheinlichkeit von 20% zu Lungenkrebs führt (Wert für langjährige starke Raucher nach Evans, 1993). Dem gegenüber stehen kausale Strukturhypothesen. Diese Hypothesen behaupten das Vorliegen eines spezifischen Kausalmodells. Damit konstatiert eine kausale Strukturhypothese nicht nur, dass zwischen den beteiligten Ereignissen gewisse Kausalzusammenhänge vorliegen, sondern auch, dass alle strukturellen Implikationen des Modells erfüllt sind. Ein Beispiel für solche Strukturhypothesen bieten die verschiedenen Modelle zur Entstehung von Magengeschwüren, die im Laufe der achtziger Jahre vorgeschlagen wurden (vgl. Thagard, 1999). Eines der Modelle sah die bakteriologische Infektion des Magens als eine Folge der Magenübersäuerung an, welche auch zu Geschwüren führe. Das andere Modell behauptete, dass die Bakterien die gemeinsame Ursache von Übersäuerung und Magengeschwüren sind.

Im folgenden Teilkapitel wird es zunächst um die Prüfung von kausalen Zusammenhangshypothesen gehen. Dabei werden wir sehen, dass eine solche Prüfung durch die Bestimmung der Kausalzusammenhänge erfolgt. Danach werden wir uns der Prüfung von kausalen Strukturhypothesen, d.h. der Prüfung von Kausalmodellen zuwenden. Diese verlangt neben der Bestimmung der Kausalzusammenhänge auch eine Überprüfung der strukturellen Implikationen.

2.3.1 Prüfen von Zusammenhangshypothesen

Beim Lernen eines Kausalzusammenhangs ist es noch offen, ob und wenn ja, was für eine Beziehung zwischen den Ereignissen vorliegt. Beim Hypothesenprüfen wird dagegen von einer speziellen Vermutung ausgegangen. Trotz dieses Unterschieds ist das normativ korrekte Vorgehen beim Lernen wie beim Hypothesenprüfen gleich. In beiden Fällen gilt es den Kausalzusammenhang zu bestimmen. Daher ist auch beim Hypothesenprüfen die

angenommene Richtung des Kausalzusammenhangs von entscheidender Bedeutung. Die Kontingenz oder Power muss von der Ursache hin zur Wirkung bestimmt werden. Können zur Überprüfung der Hypothesen zusätzlich Experimente durchgeführt werden, dann sollte die mutmaßliche Ursache manipuliert werden. Nur die Veränderung der Ursache führt zu einer Veränderung des Effektes, eine Manipulation des Effektes lässt die Ursache unverändert.

Empirische Befunde

In Untersuchungen zum wissenschaftlichen Denken wurde u.a. auch das Prüfen von Hypothesen untersucht. Die Frage dabei war, welchen Einfluss spezifische Vorerwartungen (sprich Hypothesen) auf die Bestimmung von Kausaleinflüssen hatten. Dabei zeigten sich folgende Effekte: War die Erwartung, dass zwischen der Ursache und dem Effekt ein positiver Zusammenhang vorliegt, so wurde zwar ein positiver Zusammenhang auch gut erkannt, aber ein Nullzusammenhang, welcher die Hypothese widerlegte, wurde nicht entdeckt (Chapman & Chapman, 1969; Kuhn et al., 1995). Die Probanden nahmen einen Zusammenhang wahr, obwohl objektiv gesehen kein statistischer Zusammenhang vorhanden war. Dieser Effekt war besonders stark bei Aufgabenstellungen im sozialen Bereich (Kuhn et al., 1995; Schauble, 1996). War dagegen die Hypothese, dass zwischen den Ereignissen kein Zusammenhang gegeben ist, dann wurde sowohl ein Nullzusammenhang als auch ein positiver Zusammenhang richtig erkannt (Kuhn et al., 1995).

Eine häufig untersuchte Fragestellung war, ob Probanden in der Lage sind, Hypothesen in Form von Wenn-Dann-Aussagen richtig zu überprüfen. Das bekannteste Untersuchungsparadigma in diesem Zusammenhang ist die sogenannte Wason Selection Task (Wason, 1960, 1968). Die Aufgabe ist herauszufinden, ob eine Aussage der Form ‚Wenn U dann E‘ wahr ist. Ein bekanntes Beispiel ist die Regel: ‚Wenn auf der Vorderseite ein Vokal (U) ist, dann ist auf der Rückseite ein gerade Zahl (E).‘²⁸ Zur Überprüfung der Regel lag den Probanden je ein Fall von U, ~U, E und ~E vor, die sie untersuchen durften. Da die Implikation $U \rightarrow E$ äquivalent zu $\sim(U \cdot \sim E)$ ist, sollten die Probanden überprüfen, ob bei den gezeigten Fällen die Kombination U·~E gegeben ist. D.h., sie sollten die Fälle U und ~E näher untersuchen.

²⁸ Kausalhypothesen lassen sich auch als Regeln auffassen. Die Hypothese, dass das Ereignis U das Ereignis E hervorruft, entspricht der Regel ‚Wenn die Ursache U vorliegt, dann tritt auch der Effekt E auf.‘

Die gleiche Vorhersage ergibt sich aus der Anwendung syllogistischer Schluss schemata. Wenn die Regel $U \rightarrow E$ gilt und U vorliegt, kann über den modus ponens gefolgert werden, dass E gegeben sein muss,

$$(U \rightarrow E).U \Rightarrow E \quad (\text{modus ponens}).$$

Ebenso folgt aus dem modus tollens, dass bei Gültigkeit der Regel aus $\sim E$ $\sim U$ folgt.

$$(U \rightarrow E).\sim E \Rightarrow \sim U \quad (\text{modus tollens}).$$

Aus den beiden anderen Fällen ($\sim U$ und E) kann nichts gefolgert werden oder anders gesagt: Die Gültigkeit der Regel kann von diesen Fällen nicht widerlegt werden. Das Ergebnis in einer Vielzahl von Studien war aber, dass die Fälle U und $\sim E$ nur von einer Minderheit der Probanden berücksichtigt wurden. Weitaus häufiger wurden die Fälle U und E untersucht. Diese Prüfstrategie wurde als der (fehlgehende) Versuch interpretiert, die Hypothese zu bestätigen. Da sich diese Tendenz auch in anderen Experimenten zeigte (Wason, 1968), wurde lange Zeit von einem generellen ‚confirmation bias‘ beim Hypothesenprüfen ausgegangen.

Der Annahme eines generellen confirmation bias wurde in den letzten Jahren sowohl theoretisch wie empirisch widersprochen. So konnte gezeigt werden, dass Probanden, sofern es sich um soziale Sachverhalte handelt, durchaus in der Lage sind, solche Regeln normativ richtig zu überprüfen (Cheng & Holyoak, 1986; Gigerenzer & Hug, 1992; Johnson-Laird et al., 1972). Ein Beispiel für eine solche Regel ist: „Wenn jemand unter 18 ist, dann darf er keinen Alkohol trinken“. Die Probanden untersuchten richtigerweise die Person unter 18 und die Person, die Alkohol trank. Theoretisch wurde gegen den confirmation bias argumentiert, dass es sich nicht um einen bias, also ein fehlerhaftes Vorgehen, sondern um eine positive Teststrategie handelt. Bei dieser Strategie wird aus einer Hypothese ein Datum vorhergesagt und dann überprüft, ob dieses Datum gegeben ist. Liegt das Datum nicht vor, dann muss die Hypothese falsch sein. D.h., auch eine positive Teststrategie kann zu einer Falsifikation der Hypothese führen (Klayman & Ha, 1987). Angewandt auf Kausalzusammenhänge bedeutet eine solche Strategie, dass bei Vorliegen der Ursache untersucht wird, ob der Effekt eintritt und dass bei Vorliegen des Effektes geprüft wird, ob die Ursache zuvor aufgetreten ist. Dagegen wird bei einer negativen Teststrategie vorhergesagt, was nicht beobachtet werden dürfte und dann überprüft, ob dies auch nicht gegeben ist. Es konnte gezeigt werden (Klayman & Ha, 1987), dass die Wahrscheinlichkeit, eine Hypothese richtig zu falsifizieren, bei einer positiven Teststrategie sogar größer ist als bei einer negativen, sofern die a priori Wahrscheinlichkeit des vorgesagten Datums gering ist, $P(\text{Datum}) < 0,5$.²⁹ Dies ist unter normalen Bedingungen sehr häufig der Fall. Es ist in der

²⁹ Für Details siehe Klayman & Ha, 1987, S. 216-218.

Regel sehr unwahrscheinlich, dass ein bestimmtes Ereignis zu einem bestimmten Zeitpunkt per Zufall auftritt. Zum Beispiel sollte bei der Überprüfung der Hypothese, dass BSE (Bovine Spongiforme Encephalitis) durch Prionen (infektiöse Proteine) hervorgerufen wird, überprüft werden, ob bei BSE kranken Tieren Prionen nachweisbar sind (positive Teststrategie), da die Wahrscheinlichkeit, überhaupt Prionen zu finden, insgesamt sehr gering ist.³⁰ Werden diese nun bei erkrankten Tieren in großer Zahl nachgewiesen, so ist dies ein überzeugender Beweis für die Prionenhypothese. Zu zeigen, dass alle gesunden Rinder keine Prionen aufweisen, scheint ein wenig sinnvolles Vorgehen zu sein.³¹ Eine negative Teststrategie ist beispielsweise bei Herpes angezeigt. Zum Nachweis, dass Herpesviren die bekannten Bläschen hervorruft, ist es keine gute Strategie, die Personen mit Bläschen auf Viren zu untersuchen. Wir alle haben mit hoher Wahrscheinlichkeit eine Herpesinfektion. Hier gilt es deshalb nachzuweisen, dass nicht Infizierte niemals Bläschen bekommen.

Neuere Untersuchungen zur Prüfung wissenschaftlicher Hypothesen haben gezeigt, dass es eine starke Tendenz zu einer positiven Teststrategie gibt (Klahr, 2000; Koslowski, 1996). So sollten beispielsweise die Probanden in den Experimenten von Klahr, Dunbar und Mitarbeitern (Klahr, 2000) herausfinden, wie ein bestimmter Steuerungsbefehl („repeat X“) das Verhalten eines Roboters beeinflusst. Zur Untersuchung der Wirkung des Befehls konnten die Probanden Programme schreiben und diese vom Roboter ausführen lassen. Die meisten Probanden hatten zunächst die Hypothese, dass „repeat X“ das gesamte Programm x mal ablaufen lässt oder den vorherigen Schritt im Programm x-mal wiederholt.³² Aus der Hypothese wurde ein Programm abgeleitet und das Verhalten des Roboters vorhergesagt (positive Teststrategie). Alternativhypothesen wurden nur selten berücksichtigt. Bei einer Vielzahl von Aufgaben unterschiedlicher Komplexität zeigte sich, dass diese positive Teststrategie erfolgreich war. So fanden bei der oben geschilderten Aufgabe 95% der Teilnehmer die richtige Lösung. Bei sehr komplexen Aufgaben waren es immerhin noch 50% (Klahr, 2000, Studien 1,2 und 9).

Es wurden auch Experimente durchgeführt, in denen die Probanden eine reale wissenschaftliche Entdeckung nachvollziehen sollten. Die Aufgabe war den Mechanismus aufzuklären, über den bestimmte Gene andere Gene aktivieren (Dunbar, 1993). Dazu wurde

³⁰ Derzeit lassen sich meines Wissens Prionen nur in sehr hohen Konzentrationen nachweisen, d.h. in nur einem fortgeschrittenen Stadium der Erkrankung.

³¹ Das Raben-Paradox (Hempel, 1965) liefert eine weitere Begründung für den Einsatz einer positiven Teststrategie. Das Paradox lautet wie folgt: Die Hypothese, dass alle Schwäne weiß sind, ist logisch äquivalent zu der Aussage, dass alles was nicht weiß ist, kein Schwan ist. Damit wird aber die Beobachtung eines schwarzen Raben zu einem Beleg für die Hypothese. Logisch ist diese Schlussfolgerung einwandfrei, dennoch scheint sie falsch zu sein. Daraus, dass gesunde Schweine keine Prionen haben, lässt sich nicht folgern, dass BSE durch Prionen ausgelöst wird.

³² Beide Hypothesen waren aber falsch. „Repeat X“ wiederholte das Programm ab dem x-ten Programmbefehl. Die Ausgangshypothese wurde daher auch bei einer positiven Teststrategie sehr schnell widerlegt.

den Probanden die Ausgangshypothese vermittelt, dass bestimmte Gene andere aktivieren können.³³ In den Daten war auch ein positiver Zusammenhang zwischen den Genen A und B gegeben. Tatsächlich lagen dieser vermeintlich positiven Wirkung aber zwei inhibitorische Mechanismen zu Grunde. Das auslösende Gen A hemmte ein weiteres Gen C, welches wiederum das Gen B hemmte. Wurde also das Gen A aktiv, dann wurde das Gen B enthemmt und dadurch aktiviert. Nur sehr wenige Probanden waren in der Lage, diesen Kontrollmechanismus aufzuklären. Die Schwierigkeit bei dieser Aufgabe lag darin, dass ein vollkommen neues Konzept entwickelt werden musste, Aktivierung durch Hemmung einer Hemmung. Diese Anforderung ging deutlich darüber hinaus, einen vermutlich positiven Zusammenhang zu überprüfen.³⁴

Bereits im Teilkapitel über das Lernen von Kausalzusammenhängen wurden die Studien erwähnt, die sich mit der Frage beschäftigt haben, welche Evidenzen die möglichen Beobachtungen für eine Zusammenhangshypothese darstellen. Das gemeinsame Auftreten der Ursache und des Effektes erwies sich stets als stark positive Evidenz. Das alleinige Auftreten nur der Ursache oder nur des Effektes wurde als negative Evidenz gewertet. Die gemeinsame Abwesenheit von Ursache und Effekt wurde entgegen normativer Modelle als entweder nicht aussagekräftig oder nur leicht positive Evidenz angesehen (Shaklee & Tucker, 1980).

Andere Studien haben sich mit der Frage beschäftigt, welchen Einfluss die Menge an bestätigender Evidenz auf die Revision einer Hypothese bei widerlegenden Daten hat (Keinath & Krems, 1999; Krems & Johnson, 1995). Die Aufgabe war, Atome in einer Black Box anhand der Ablenkung von Lichtstrahlen zu lokalisieren. Da die Atome die Lichtstrahlen auf eine den Probanden bekannte Weise ablenkten, entwickelten die Probanden sehr schnell, schon aufgrund einzelner Beobachtungen, Hypothesen über den Ort der Atome. Die den Probanden gezeigten Daten waren aber so gestaltet, dass sie grundsätzlich auch mit anderen als den von den Probanden typischerweise inferierten Hypothesen vereinbar waren. Die Befunde zeigen, dass die Probanden eine starke Tendenz hatten, ihre erste Hypothese beizubehalten. Je mehr positive Evidenzen für diese Hypothese vorlagen, desto unwahrscheinlicher war es, dass diese aufgrund widersprüchlicher Befunde aufgegeben wurde. Vielmehr wurden Zusatzannahmen gemacht, um die widersprechenden Daten zu erklären. Erst nachdem für eine Widerlegung der Ausgangshypothese ausreichend viele Evidenzen vorlagen, wurde diese zu Gunsten einer anderen Hypothese aufgegeben (Keinath & Krems, 1999).

³³ Von dieser Hypothese waren die Forscher tatsächlich ausgegangen.

³⁴ Die Entdeckung der genetischen Kontrollmechanismen wurde mit dem Nobelpreis ausgezeichnet.

Zusammenfassend lässt sich sagen, dass Personen generell gut in der Lage sind, kausale Zusammenhangshypothesen zu testen, insbesondere wenn sie selbst experimentieren dürfen oder schlüssige Daten vorgelegt bekommen. Sie berücksichtigen dabei die in den Daten vorhandenen Kontingenzen zwischen den Ereignissen.

Bei der Prüfung setzen sie bevorzugt eine positive Teststrategie ein. Die Evidenzen verrechnen sie dann häufig über vereinfachende Heuristiken, v.a. die Differenzbildung aus bestätigenden und widerlegenden Evidenzen. Diese Heuristiken haben sich als sehr erfolgreich erwiesen, führten sie doch in der Mehrheit der Fälle zur Entdeckung des tatsächlichen Zusammenhangs.

2.3.2 Prüfen von Kausalzusammenhängen innerhalb eines Kausalmodells

Häufig sollen nicht nur einzelne isolierte Kausalhypothesen, sondern mehrere Hypothesen über miteinander verbundene Kausalzusammenhänge überprüft werden. In diesem Fall ist es wie beim Lernen entscheidend, in welches Kausalmodell diese Zusammenhänge eingebettet sind. Gehen wir wieder von den drei grundlegenden Modellen aus. Bilden die Zusammenhänge ein Gemeinsame-Ursache Modell, dann ist die Wirkung der Ursache auf den einen Effekt unabhängig von der Wirkung auf den anderen Effekt. Deshalb ist es legitim, die zugehörigen Kausalhypothesen einzeln, unabhängig voneinander zu überprüfen. In diesem Fall kann das im vorigen Abschnitt dargestellte Vorgehen angewandt werden. Liegt ein Kettenmodell vor, dann ist unter der Markovbedingung die Wirkung der ersten Ursache auf das intermittierende Ereignis unabhängig von der Wirkung des mittleren Ereignisses auf den letzten Effekt und umgekehrt. Daher können auch hier die beiden Hypothesen einzeln getestet werden. Beim Gemeinsamer-Effekt Modell trifft dies aber nicht zu. Hier wirken mehrere Ursachen auf denselben Effekt ein. Bei der Überprüfung der Hypothese, dass die eine Ursache den Effekt hervorruft, muss daher die Alternativhypothese, dass der Effekt durch die andere Ursache hervorgerufen wird, mit berücksichtigt werden. Es reicht nicht aus, die Hypothesen einzeln, unabhängig voneinander zu überprüfen. Wie beim Lernen mehrerer Kausalzusammenhänge muss daher zur Bestimmung des einen Kausalzusammenhangs der Einfluss der anderen Ursache kontrolliert werden. In diesem Fall ist bedingte Kontingenz bzw. die bedingte Power der richtige Indikator für einen Kausalzusammenhang. Sehen wir uns die möglichen Hypothesenkonstellationen beim Gemeinsamer-Effekt Modell im Einzelnen an. Ist die Hypothese, dass beide Ursachen positiv auf den Effekt wirken, so muss bei der Bestimmung der Kontingenz der Zielursache auf den Effekt die alternative Ursache konstant abwesend

gehalten werden. Nur so kann sichergestellt werden, dass die beobachtete bedingte Kontingenz nicht durch eine Konfundierung mit dieser anderen Ursache bedingt ist. Ist die Hypothese, dass die Zielursache positiv auf den Effekt wirkt, aber die Alternativursache den Effekt verhindert, so muss ebenfalls die Alternativursache konstant abwesend gehalten werden. Ansonsten würde die Kontingenz die Wirkung der Zielursache unterschätzen. Eine dritte mögliche Hypothesenkonstellation ist, dass die Zielursache den Effekt verhindert, die andere Ursache ihn aber hervorruft. Damit die negative Wirkung der Ursache in diesem Fall überhaupt sichtbar wird, muss der Effekt zumindest in der Kontrollbedingung auftreten ($P(e|\sim c) > 0$). Daher muss mindestens eine alternative generative Ursache konstant anwesend gehalten werden. Eine weitere Möglichkeit ist, dass es sich bei der alternativen Ursache um eine notwendige Randbedingung (enabling condition) handelt. Enabling conditions sind Bedingungen, die vorliegen müssen, damit eine andere Ursache wirken kann, die aber selbst nicht auf den Effekt wirken. So ist Gravitation eine notwendige Bedingung dafür, dass ein Weinglas zu Boden fällt, sie ist aber nicht die Ursache für den Rotweinfleck im Teppich. Enabling conditions müssen natürlich konstant anwesend gehalten werden.³⁵

Empirische Befunde

In mehreren Experimenten wurde untersucht, ob Personen bei einem Gemeinsamer-Effekt Modell das Vorzeichen der zu überprüfenden Hypothese berücksichtigen. Dazu wurden Probanden gebeten die Hypothese zu prüfen, dass eine Ursache (Medikament) zu einem Effekt (Schwangerschaft bei Katzen) führt (Wu & Cheng, 1999). Das Medikament selbst wies in allen Bedingungen keinen statistischen Zusammenhang mit dem in Frage stehenden Effekt auf. Manipuliert wurde die Wahrscheinlichkeit des Effektes ohne Vorliegen der Ursache ($P(e|\sim u)$). In einer der beiden extremen Ausprägungen war der Effekt nie vorhanden, wenn das Medikament nicht verabreicht wurde, ($P(e|\sim u) = 0$), in der anderen Bedingung trat der Effekt unabhängig von der Ursache immer auf ($P(e|\sim u) = 1$). Wenn die Probanden das Vorzeichen der zu überprüfenden Hypothese berücksichtigen, dann sollten sie in den beiden Datenbedingungen trotz der identischen Nullkontingenz der Ursache zu unterschiedlichen Urteilen gelangen. Da die Zielursache laut Hypothese den Effekt hervorruft, sollten alternative generative Ursachen möglichst konstant abwesend gehalten werden. Dies ist offensichtlich in der Bedingung, in welcher der Effekt immer eintritt, nicht

³⁵ Bei realen Sachverhalten ist die Anzahl der Ereignisse, die einen Einfluss auf den Zieleffekt haben, sehr groß. Das Gemeinsamer-Effekt Modell umfasst viele Ursachen und notwendige Randbedingungen (enabling conditions). Durch das Konstanthalten der anderen Ursachen wird die Menge der Fälle begrenzt, die für die Bestimmung eines Kausalzusammenhangs zu untersuchen sind. Ist die Hypothese, dass die Zielursache positiv auf den Effekt wirkt, dann müssen z.B. nur die Fälle betrachtet werden, in denen alle alternativen Ursachen abwesend und die enabling conditions anwesend sind.

gelingen. Es sind stets andere Ursachen vorhanden, welche den Effekt verursachen. Daher kann in diesem Fall die Wirkung des Medikaments nicht bestimmt werden. Ganz anders aber, wenn der Effekt bei Abwesenheit der Ursache nie auftritt. Offensichtlich sind hier keine anderen generativen Ursachen vorhanden. Die Kontingenz gibt also direkt die Stärke des Kausalzusammenhangs an. Die Urteile der Probanden bestätigten die Vorhersagen. Trat der Effekt immer ein, so gaben sie an, keine Aussage machen zu können, blieb er ohne Ursache stets aus, so gaben sie an, dass die Ursache keine Wirkung habe (Wu & Cheng, 1999). Dieselbe Untersuchung wurde auch mit der Hypothese durchgeführt, dass die in Frage stehende Ursache den Effekt hemmt. In diesem Fall kehren sich die Vorhersagen um. Alternative generative Ursachen müssen in diesem Fall konstant anwesend gehalten werden, damit die hemmende Wirkung der Ursache sichtbar werden kann. Tritt der Effekt also in Abwesenheit der Ursache immer auf, dann sind dies ideale Bedingungen für eine Prüfung der Hypothese. Dagegen kann die verhütende Wirkung des Medikaments nicht untersucht werden, wenn die zu unterdrückende Schwangerschaft bei den Versuchstieren nie auftritt. Die Probanden zeigten sich auch hier sensitiv für die unterschiedlichen Bedingungen. Trat der Effekt ohne Ursache immer auf, so schlossen sie richtig, dass das Medikament nicht wirkt. Blieb der Effekt stets aus, so urteilten sie, dass keine Aussage über die präventive Wirkung des Medikaments möglich sei (Wu & Cheng, 1999). Diese Ergebnisse belegen, dass Personen das Vorzeichen der zu überprüfenden Hypothese und die vorliegenden Daten normativ richtig berücksichtigen.³⁶

Insgesamt sind die Befunde für eine normativ richtige Nutzung von Kausalmodellen beim Hypothesenprüfen noch dürftig. Nimmt man aber die Befunde der Untersuchungen zur Bestimmung von Kausalzusammenhängen mit hinzu, dann ist davon auszugehen, dass Probanden durchaus in der Lage sind, Hypothesen über Kausalzusammenhänge innerhalb eines Modells korrekt zu prüfen.

³⁶ Aus der Definition der kausalen Power folgen dieselben Vorhersagen. So ist die generative Power definiert als $p_g = \Delta P / 1 - P(e|\sim c)$, d.h. sie kann nur bestimmt werden, wenn $P(e|\sim c) < 1$. Die präventive Power ist definiert als $p_i = -\Delta P / P(e|\sim c)$. Daraus ergibt sich als Randbedingung, dass $P(e|\sim c) > 0$ sein muss. Daher sprechen die Ergebnisse auch für eine Verwendung von kausaler Power als Maß für Kausalität.

2.3.3 Prüfen von kausalen Strukturhypothesen

In den beiden vorangegangenen Abschnitten haben wir uns mit dem Prüfen von Hypothesen über einzelne Kausalzusammenhänge beschäftigt. Bei dieser Aufgabe sind die hinter den Zusammenhängen stehenden Kausalmodelle relevant. Doch auch die angenommenen Kausalmodelle selbst sind häufig nur Hypothesen, Hypothesen über das Vorliegen einer bestimmten Struktur. Struktur- und Zusammenhangshypothesen unterscheiden sich grundsätzlich voneinander. Eine Zusammenhangshypothese besagt, dass ein bestimmter Kausalzusammenhang vorliegt. Sie ist widerlegt, wenn eine Bestimmung dieses Zusammenhangs ergibt, dass dieser nicht vorliegt, ein anderes Vorzeichen hat oder nicht die vermutete Stärke aufweist. Eine Strukturhypothese besagt, dass die an einem Modell beteiligten Zusammenhänge eine bestimmte Struktur bilden. Sie ist widerlegt, wenn die Implikationen dieser Struktur nicht erfüllt sind. Ein Beispiel soll diesen Unterschied veranschaulichen.

Beispiel für die Unterscheidung von kausalen Zusammenhangs- und kausalen Strukturhypothesen

Viele Frauen nehmen Hormone ein, um eine ungewollte Schwangerschaft zu verhindern. Stellen wir uns vor, ein neues Präparat würde auf den Markt kommen. Eine Hypothese in Bezug auf das Präparat ist, dass seine Einnahme Schwangerschaften verhindert, eine zweite Vermutung ist, dass die Hormone auch Thrombosen vermeiden helfen. Die beiden vermuteten Kausalzusammenhänge bilden also ein Gemeinsame-Ursache Modell.

Prüfen wir zunächst aber die beiden Zusammenhangshypothesen. Wie wir oben gesehen haben, können bei einem Gemeinsame-Ursache Modell beide Hypothesen unabhängig voneinander geprüft werden. Die Abbildung auf der folgenden Seite gibt ein fiktives Zahlenbeispiel.

	Thrombose		Keine Thrombose	
	Schwangerschaft	Keine Schwangerschaft	Schwangerschaft	Keine Schwangerschaft
Einnahme des Präparats	5	10	0	20
Keine Einnahme des Präparat	35	0	10	20
Kontingenzen	$\Delta P_{\text{Präparat-Thrombose}} = 15/35 - 35/65 = -0,11$ $\Delta P_{\text{Präparat-Schwangerschaft}} = 5/35 - 45/65 = -0,55$ $\Delta P_{\text{Schwangerschaft-Thrombose}} = 40/50 - 10/50 = +0,60$ $\Delta P_{\text{Schwangerschaft-Thrombose Präparat}} = 5/5 - 10/30 = +0,67$ $\Delta P_{\text{Schwangerschaft-Thrombose ~Präparat}} = 35/45 - 0/20 = +0,77$			

Abb. 2.12: Fiktives Zahlenbeispiel zur Unterscheidung von Zusammenhangs- und Strukturhypothesen. Erläuterungen siehe Text

Eine Überprüfung der beiden Hypothesen ergibt, dass das Präparat die Wahrscheinlichkeit einer Schwangerschaft deutlich senkt ($\Delta P = -0,55$) und auch Thrombosen zumindest leicht verhindert ($\Delta P = -0,11$). Es sollten also beide Zusammenhangshypothesen angenommen werden.

Wie sieht es aber mit der Strukturhypothese aus? Bilden die drei Ereignisse wirklich ein Gemeinsame-Ursache Modell? Wenn ein Gemeinsame-Ursache Modell tatsächlich vorliegt, dann sollten auch die strukturellen Implikationen für dieses erfüllt sein. Wenn das Präparat sich sowohl auf eine Schwangerschaft als auch auf Thrombosen auswirkt, dann sollten diese beiden Ereignisse miteinander korrelieren. Genau das ist auch der Fall ($\Delta P = +0,60$). Wenn der Zusammenhang zwischen Thrombosen und Schwangerschaft nur auf die Wirkung des Präparates zurückzuführen ist, dann sollte der Zusammenhang verschwinden, wenn diese Ursache konstant gehalten wird. Oder anders gesagt: Die Effekte einer gemeinsamen Ursache sind konditional auf diese Ursache unabhängig voneinander. In dem vorliegenden Beispiel trifft dies aber nicht zu. Der Zusammenhang ist sowohl in der Gruppe der Frauen, die das Präparat nehmen, als auch bei denen, die es nicht nehmen, immer noch sehr hoch. Daher muss das vermutete Gemeinsame-Ursache Modell falsch sein. Die hohe Korrelation von Schwangerschaft und Thrombosen kann nicht durch dieses Modell erklärt werden. Daraus folgt, dass die Strukturhypothese abgelehnt werden muss.

In diesem Beispiel werden also beide Zusammenhangshypothesen angenommen, die Strukturhypothese wird aber zurückgewiesen. Wie kann es zu einem solchen Fall kommen? Beim Prüfen der Zusammenhangshypothesen sind wir davon ausgegangen, dass das Modell

stimmt und haben auf dieser Grundlage die Kausalzusammenhänge bestimmt. Stellt sich nun heraus, dass die Strukturhypothese falsch war, ist damit auch die Prüfung der Zusammenhangshypothesen falsch. Der starke Zusammenhang zwischen Schwangerschaft und Thrombose spricht dafür, dass zwischen diesen ein weiterer Kausalzusammenhang bestehen muss. Da Thrombosen bekanntermaßen nicht zu Schwangerschaften führen, muss ein Zusammenhang in der anderen Richtung bestehen. Das ist auch tatsächlich der Fall. Das korrekte Modell besteht aus drei Kausalzusammenhängen. Das Präparat hat einen Einfluss auf Schwangerschaften und auf die Entstehung von Thrombosen. Zusätzlich erhöht eine Schwangerschaft das Thromboserisiko. Wenn eine Schwangerschaft aber eine weitere Ursache von Thrombosen ist, dann muss diese bei der Bestimmung des Kausalzusammenhangs zwischen dem neuen Präparat und Thrombosen berücksichtigt werden. Bestimmt man nun folgerichtig die bedingte Kontingenz, so zeigt sich, dass bei den Frauen, die nicht schwanger sind, das Präparat die Wahrscheinlichkeit von Thrombosen erhöht ($\Delta P_{\text{Präparat-Thrombose}|\text{-Schwangerschaft}} = 10/30 - 0/20 = +0,33$)! Eine Bestimmung des Kausalzusammenhangs zwischen dem Präparat und Thrombosen auf Grundlage des erweiterten Kausalmodell ergibt also, dass die Hypothese, dass das Präparat Thrombosen verhindert, falsch ist. Der negative statistische Zusammenhang zwischen beiden kommt dadurch zustande, dass das Präparat Schwangerschaften verhindert, welche ein sehr hohes Thromboserisiko bergen.

Fassen wir noch einmal die wichtigsten Punkte aus dem Beispiel zusammen. Beim Prüfen von Zusammenhangshypothesen wird die Gültigkeit des Kausalmodells vorausgesetzt, beim Prüfen von Strukturhypothesen wird diese Annahme überprüft. Bei Ersterem reicht es aus, die normativ richtigen, für Kausalität indikativen statistischen Zusammenhänge zu bestimmen. Beim Prüfen von Strukturen müssen deren implizierte bedingte und unbedingte Abhängigkeiten bestimmt werden. Nur wenn diese vorliegen, kann die Strukturhypothese zutreffen.

Normativ richtiges Vorgehen beim Prüfen von kausalen Strukturhypothesen

Bei der Überprüfung von Kausalmodellen müssen sowohl die Kausalzusammenhänge als auch die strukturellen Implikationen der Modelle geprüft werden. Wie wir im ersten Kapitel gesehen haben, unterscheiden sich die strukturellen Implikationen verschiedener Kausalmodelle. Abbildung 2.13 zeigt nochmals die drei Grundmodelle.

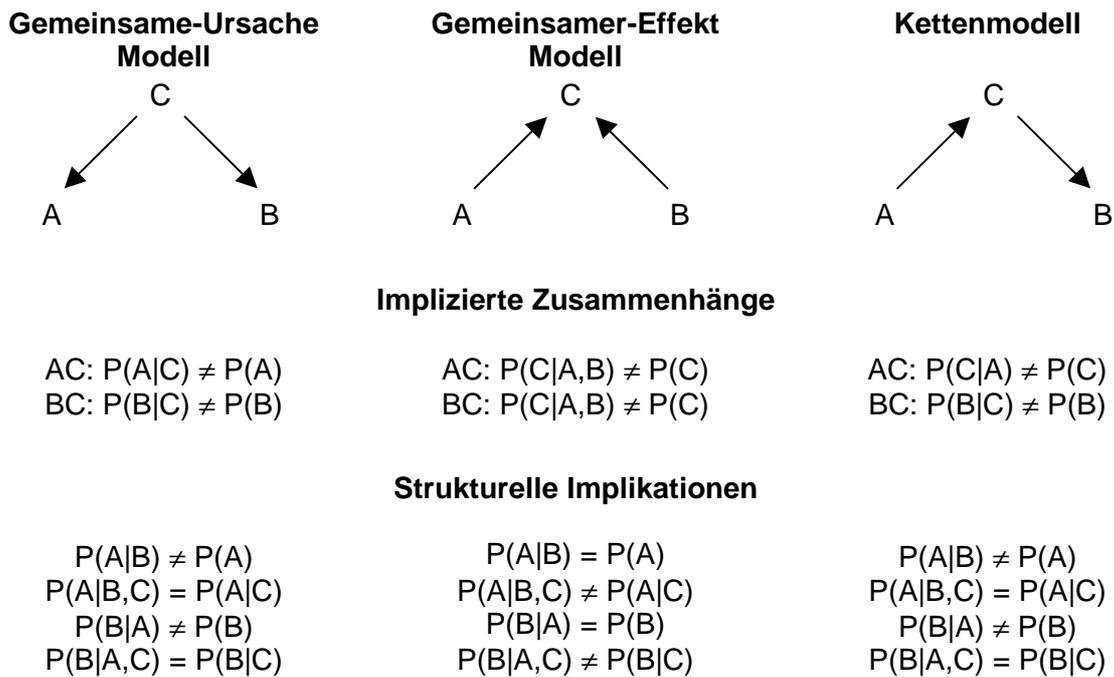


Abb. 2.13: Zu prüfende Abhängigkeiten bei den grundlegenden Kausalmodellen

Ein Kausalmodell trifft zu, wenn erstens eine Abhängigkeit zwischen den Ursachen und ihren direkten Effekten besteht und zweitens zusätzlich die strukturellen Implikationen erfüllt sind. Ist die Hypothese, dass das Ereignis U die beiden Ereignisse A und B hervorruft (Gemeinsame-Ursache Modell), dann muss zuerst überprüft werden, ob eine Abhängigkeit zwischen den Ereignissen C und A und den Ereignissen C und B besteht. Dies kann auch über die Bestimmung der Stärke der beiden vermuteten Kausalzusammenhänge erfolgen. Liegt eine Abhängigkeit vor bzw. wirkt die Ursache auf die beiden anderen Ereignisse, dann ist als nächstes zu überprüfen, ob die beiden vermeintlichen Effekte kovariieren.³⁷ Die Effekte sollten aber unabhängig werden, wenn ihre gemeinsame Ursache konstant gehalten wird. Nur wenn alle diese Überprüfungen zu einem positiven Ergebnis führen, ist die Strukturhypothese mit den Daten vereinbar.

Ist die Hypothese, dass das Ereignis A das Ereignis C und dieses wiederum das Ereignis B verursacht (Kettenmodell), dann muss ebenfalls als Erstes überprüft werden, ob ein Zusammenhang zwischen den direkt kausal verbundenen Ereignissen besteht. Ist dies der Fall, dann muss zusätzlich geprüft werden, ob der letzte Effekt in der Kette mit der ersten Ursache kovariiert. Dieser Zusammenhang sollte aber verschwinden, wenn das vermittelnde Ereignis C konstant gehalten wird. Ist dies der Fall, dann stimmen die Daten mit dem vermuteten Modell überein. Wie aus diesen Ausführungen und der Abbildung klar wird, haben das Gemeinsame-Ursache Modell und das Kettenmodell die gleichen Implikationen. D.h., diese beiden Modelle können nur aufgrund von Daten nicht unterschieden werden.

Auch wenn alle implizierten Abhängigkeiten vorliegen, könnten beide Hypothesen zutreffen. Es werden also weitere Kriterien wie die zeitliche Abfolge oder die experimentelle Manipulation der einzelnen Ereignisse benötigt, um zwischen diesen zu unterscheiden.

Das Gemeinsamer-Effekt Modell hat dagegen spezifische, von allen anderen Modellen verschiedene, strukturelle Implikationen. Bei diesem Modell muss wiederum als Erstes überprüft werden, ob Zusammenhänge zwischen den beiden Ursachen und dem Effekt vorliegen. Sofern es keine gegenteiligen begründeten Annahmen gibt, sollten die beiden vermeintlichen Ursachen unabhängig voneinander sein. Wird ihr gemeinsamer Effekt aber konstant gehalten, dann sollte sich eine negative bedingte Abhängigkeit zeigen. Sind diese Implikationen gegeben, dann trifft die Gemeinsamer-Effekt Hypothese zu. Diese Hypothese kann also eindeutig bestätigt oder widerlegt werden.

Fassen wir die obigen Überlegungen zum Prüfen von Strukturhypothesen zusammen, dann gilt folgendes: Das Vorliegen der vermuteten Kausalzusammenhänge ist notwendig dafür, dass eine Strukturhypothese zutrifft. Die Erfüllung der strukturellen Implikationen ist beim Gemeinsame-Ursache Modell und beim Kettenmodell notwendig und beim Gemeinsamer-Effekt Modell hinreichend für eine Bestätigung der Hypothese.

Empirische Evidenzen

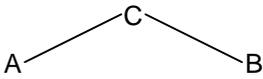
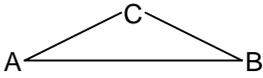
Meines Wissens sind bisher keine Experimente zur Prüfung kausaler Strukturhypothesen durchgeführt worden. In den bisher durchgeführten Studien war der verwendete Sachverhalt entweder so gestaltet, dass das Kausalmodell aufgrund des Vorwissens eindeutig bestimmt war, oder das Modell wurde vorgegeben. Die Aufgabe der Probanden war es lediglich, die einzelnen Kausalzusammenhänge innerhalb des Modells zu überprüfen. Daher geben diese Studien keinen Aufschluss darüber, ob Personen in der Lage sind, normativ richtig kausale Strukturhypothesen zu testen.

Offene Fragen

Da bisher keine Studien zur Prüfung kausaler Strukturhypothesen vorliegen, wurden im Rahmen dieser Arbeit eine Reihe von Experimenten zu dieser Fragestellung durchgeführt. Die Probanden erhielten die Aufgabe anhand von Daten zu prüfen, ob ein bestimmtes Kausalmodell gegeben ist. Die Experimente haben sich dabei auf das Gemeinsame-Ursache und das Gemeinsamer-Effekt Modell beschränkt. Diese beiden Strukturmodelle vermuten Kausalzusammenhänge zwischen denselben Ereignissen, weisen beide ein zentrales Ereignis auf, das einen anderen kausalen Status als die beiden anderen Ereignisse hat, und sie haben entgegengesetzte strukturelle Implikationen. Um

³⁷ Die Höhe der Kovariation kann dabei direkt aus der Stärke der beteiligten Kausalzusammenhänge vorhergesagt werden. Siehe Kapitel 1.3.

herauszufinden, ob die Probanden die strukturellen Implikationen als Prüfkriterium heranziehen, wurden die gezeigten Daten speziell konstruiert. Ein Datensatz erfüllte die strukturellen Implikationen des Gemeinsame-Ursache Modells, der andere Datensatz die des Gemeinsamer-Effekt Modells. Die vermuteten Kausalzusammenhänge lagen in beiden Datensätzen vor. Die unten stehende Abbildung zeigt den grundlegenden Aufbau der Experimente.

	Gemeinsame-Ursache Modell $A \leftarrow C \rightarrow B$	Gemeinsamer-Effekt Modell $A \rightarrow C \leftarrow B$
Daten pro Gemeinsamer-Effekt Hypothese  $P(A B) = P(A)$ $P(A B,C) \neq P(A C)$	<i>Ablehnung</i>	<i>Annahme</i>
Daten pro Gemeinsame-Ursache Hypothese  $P(A B) \neq P(A)$ $P(A B,C) = P(A C)$	<i>Annahme</i>	<i>Ablehnung</i>

2.14: Experiment zum Prüfen kausaler Strukturhypothesen, die Einträge in den Zellen geben die normativ richtigen Prüfurteile an.

Wenn die Probanden die strukturellen Implikationen der Kausalmodelle beim Prüfen der Strukturhypothese nutzen, dann sollten sie die Gemeinsamer-Effekt Hypothese bei den gleichnamigen Daten annehmen. In diesem Fall liegen sowohl die vermuteten Kausalzusammenhänge als auch die zu diesem Modell zugehörigen strukturellen Implikationen vor. Beim Gemeinsame-Ursache Modell sind zwar die vermuteten Kausalzusammenhänge gegeben, aber die Daten widersprechen den Implikationen des Modells. Das Gemeinsame-Ursache Modell sollte aber bei den Daten ‚Pro Gemeinsame-Ursache‘ angenommen werden. Bei diesen Daten sind die strukturellen Implikationen dieses Modells erfüllt, aber die des Gemeinsamer-Effekt Modells widerlegt. Dieses Modell sollte also nun abgelehnt werden. Die einzelnen Experimente und die Befunde werden in Kapitel 3.2 vorgestellt und ausführlich diskutiert.

2.4 Generieren von Kausalhypothesen

In den vorherigen Teilkapiteln waren wir davon ausgegangen, dass bereits eine Hypothese über einen Kausalzusammenhang oder ein Kausalmodell vorliegt. Auch bei den Lernaufgaben waren zumindest die Ursache und ihr vermutlicher Effekt bekannt. Dies muss nicht immer der Fall sein. So treten teilweise Phänomene auf, für die wir noch keine Ursache kennen, beispielsweise neue Krankheitsbilder. Wenn wir nun eine Diagnose durchführen wollen, dann müssen wir zunächst Hypothesen über mögliche Ursachen entwickeln.

Daneben haben wir es auch mit unbekanntem Effekten zu tun. Ständig werden neue Ursachen in die Welt gesetzt, deren Auswirkungen nicht vollständig bekannt sind. Zu diesen zählen beispielsweise neue Medikamente und genetisch veränderte Pflanzen und Tiere. In diesem Fall gilt es mögliche Wirkungen herauszufinden. Manchmal sind sowohl die Ursache als auch der Effekt seit langem bekannt, aber ein Zusammenhang zwischen beiden wurde bisher nicht erkannt. Dies trifft vor allem bei Kausalzusammenhängen mit sehr langen Latenzzeiten zu. Beispiele hierfür sind langsam wirkende Viren („slow viruses“, Evans, 1993), die häufig erst Jahre nach einer Infektion zu einer Erkrankung führen oder fluorierte Chlorkohlenwasserstoffe (FCKWs), die Jahrzehnte nach ihrer Freisetzung zum Abbau der Ozonschicht beitragen.

Die Generierung kausaler Strukturhypothesen wird vor allem dann wichtig, wenn wir es mit Sachverhalten zu tun haben, bei denen eine Vielzahl von Faktoren relevant ist, zum Beispiel psychosomatische Erkrankungen. Viele Risikofaktoren haben einen Einfluss, das Zusammenspiel von psychischen und physiologischen Prozessen ist weitgehend unbekannt und die möglichen somatischen Auswirkungen sind vielfältig. Für eine erfolgreiche Behandlung wäre ein zutreffendes Kausalmodell der Krankheitsentstehung sehr hilfreich.

In der Wissenschaftstheorie wurden viele mögliche Wege aufgezeigt, wie neue Kausalhypothesen generiert werden können (Thagard, 1999). Dies kann per Zufall geschehen, praktische Erfordernisse können zu systematischen Untersuchungen führen oder auch Analogien helfen weiter (vgl. auch Holyoak & Thagard, 1995). Eine Diskussion der zugehörigen psychischen Prozesse würde uns aber von der Frage nach der Anwendung von Kausalmodellen wegführen. Wir werden uns daher im Folgenden auf die Rolle von Kausalmodellen bei der Generierung von Kausalhypothesen beschränken.³⁸ Der erste Abschnitt beschäftigt sich mit dem Einfluss von Kausalmodellen auf das Erlernen neuer Krankheitskonzepte, der zweite Abschnitt mit möglichen Wegen zur Ableitung von Kausalmodellen.

2.4.1 Generieren von kausalen Zusammenhangshypothesen

Mediziner werden sehr häufig mit Patienten konfrontiert, die eine Vielzahl von Symptomen gleichzeitig aufweisen und die einer Vielzahl von Risikofaktoren (=Ursachen) ausgesetzt waren. Die Frage, vor der die Medizin steht, ist, welche der Symptome zu einer Krankheit gehören und welche durch eine Mehrfacherkrankung bedingt sind. Für die Bestimmung von Krankheiten liefern Kausalmodelle wichtige Constraints. Wird eine Krankheit als die gemeinsame Ursache ihrer Symptome angesehen, dann folgt aus den strukturellen Implikationen dieses Gemeinsame-Ursache Modells, dass die Symptome untereinander kovariieren sollten. Umgekehrt ergibt sich daraus für die Bestimmung von Krankheiten, dass miteinander korrelierte Symptome zu einem Syndrom verbunden und einer Krankheit zugeordnet werden sollten. Symptome, die über verschiedene Patienten hinweg nicht miteinander kovariieren, sollten verschiedenen Krankheiten zugeordnet werden. Ganz andere Vorhersagen ergeben sich, wenn wir die Ursachen von Krankheiten betrachten. In diesem Fall ist die Krankheit der gemeinsame Effekt verschiedener Risikofaktoren. So kann eine Erkältung sowohl durch Viren wie durch Bakterien hervorgerufen werden. Die strukturellen Implikationen des Gemeinsamer-Effekt Modells besagen, dass die für die Krankheit ursächlichen Faktoren unabhängig sein sollten. Wird eine Krankheit also über ihre Ursachen bestimmt, dann sollten einzelne unabhängige Krankheitserreger zur Klassifikation herangezogen werden.³⁹

Empirische Befunde

Die Bestimmung ursachen- vs. effektbasierter Krankheitskategorien wurde meines Wissens noch nicht experimentalpsychologisch untersucht. Aber es liegen Studien zum Erlernen neuer Antwortkategorien auf der Grundlage verschiedener Kausalmodelle vor (Waldmann et al., 1995). In einem der Experimente erhielten die Probanden Informationen über das äußere Erscheinungsbild (Gewicht, Gesichtsfarbe und Transpiration) von Personen. Die Aufgabe der Probanden war es zu erlernen, welche der Personen an einer Virusinfektion litten (Gemeinsame-Ursache Bedingung) bzw., welches Erscheinungsbild bei anderen Personen zu emotionalen Reaktionen führte (Gemeinsamer-Effekt Bedingung). Die Zuordnung der gezeigten Daten zu den Antwortkategorien wurde dabei so gestaltet, dass entweder die Symptome einzeln, unabhängig voneinander die Antwort richtig vorhersagten

³⁸ So gibt es interessante Befunde zur Generierung von Hypothesen durch mehr oder minder systematisches Ausprobieren. Diese zeigen, dass Probanden in der Lage sind, aufgrund der beobachteten Veränderungen zutreffende Hypothesen zu entwickeln (Klahr, 2000, Studie 5).

³⁹ In der Medizin werden symptom- und ursachenbezogene Klassifikationssysteme nebeneinander verwendet. So werden einerseits die Erkrankungen nach den gestörten Organen systematisiert (Erkrankungen der Lunge, des Herzens etc.), andererseits durch die auslösenden Bedingungen (genetische Erkrankungen, Infektionskrankheiten etc.) (vgl. Thagard, 1999).

oder eine Korrelation zwischen den Symptomen für die Antwort prädiktiv war. Wenn die Probanden die Kausalmodelle beim Erlernen der neuen Kategorien nutzen, dann sollten sie bei einem Gemeinsame-Ursache Modell eine Korrelation der Symptome bei einer Virusinfektion erwarten und bei den entsprechenden Daten schneller die richtigen Antwortkategorien erlernen. Genau das Gegenteil ist bei einem Gemeinsamer-Effekt Modell der Fall. Bei diesem sollten die Probanden von einer Unabhängigkeit der Ursachen und ihrer Wirkung ausgehen. Sie sollten daher schneller die richtigen Antworten erlernen, wenn die einzelnen Symptome prädiktiv für die emotionale Reaktion sind. Die empirischen Befunde bestätigten diese Vorhersagen in einer Reihe von Experimenten mit unterschiedlichen inhaltlichen Aufgabenstellungen aus dem Bereich der Medizin und Physik (Waldmann et al., 1995). Die Probanden erlernten die neuen Antwortkategorien schneller, wenn die gezeigten Daten den strukturellen Implikationen der vermuteten Kausalmodelle entsprachen.

2.4.2 Generieren von kausalen Strukturhypothesen

Kausalmodelle sind von zentraler Bedeutung für eine Vielzahl von Aufgaben. Sie beeinflussen, wie Kausalzusammenhänge erlernt, Prognosen und Diagnosen gestellt und Kausalhypothesen überprüft werden. Doch wie werden diese kausalen Strukturhypothesen generiert? Im folgenden Abschnitt werden drei mögliche, normativ richtige Vorgehensweisen vorgestellt. Weitere Heuristiken, wie z.B. die Verwendung der zeitlichen Abfolge der Ereignisse, sind denkbar.

Ableitung von Kausalmodellen aus beobachteten Kovariationen von Ereignissen

Kausale Strukturmodelle können aus beobachtbaren statistischen Zusammenhängen zwischen den Ereignissen erschlossen werden. Da die verschiedenen Kausalmodelle unterschiedliche Implikationen für die Zusammenhänge zwischen den Ereignissen haben, kann durch die Bestimmung der unbedingten und bedingten Abhängigkeiten zwischen diesen Ereignissen das Kausalmodell rekonstruiert werden. Wie dies möglich ist, soll anhand eines einfachen Beispiels mit drei Ereignissen besprochen werden. Angenommen die Aufgabe sei es, die Kausalstruktur zu ermitteln, welche drei Gehirnareale A, B, C miteinander verbindet. Als Daten liegen Bilder der Durchblutung in den einzelnen Arealen vor. Wegen der schlechten zeitlichen Auflösung bei den Aufnahmen sind keine Informationen darüber verfügbar, welches Areal vor welchem anderen Areal stark durchblutet wird. Die Kausalstruktur muss also allein auf Grundlage der Kovariation der Durchblutung in den drei Arealen erschlossen werden. Die unten stehende Abbildung zeigt die vorliegenden Beobachtungen.

	Areal B stark durchblutet		Areal B schwach durchblutet	
	C stark	C schwach	C stark	C schwach
Areal A stark durchblutet	21	15	15	0
Areal A schwach durchblutet	4	10	10	25

Kontingenzen		
AB	AC	BC
$\Delta P_{AB} = 36/50 - 15/50 = 0,42$	$\Delta P_{AC} = 36/50 - 15/50 = 0,42$	$\Delta P_{BC} = 25/50 - 25/50 = 0,00$
$\Delta P_{AB C} = 15/15 - 10/35 = 0,71$	$\Delta P_{AC B} = 15/15 - 10/35 = 0,71$	$\Delta P_{BC A} = 4/14 - 10/35 = 0,00$
$\Delta P_{AB C} = 21/36 - 4/14 = 0,30$	$\Delta P_{AC B} = 21/36 - 4/14 = 0,30$	$\Delta P_{BC A} = 21/36 - 15/15 = -0,42$
BA	CA	CB
$\Delta P_{BA} = 36/50 - 15/50 = 0,42$	$\Delta P_{CA} = 36/50 - 15/50 = 0,42$	$\Delta P_{CB} = 25/50 - 25/50 = 0,00$
$\Delta P_{BA C} = 15/25 - 0/25 = 0,60$	$\Delta P_{CA B} = 15/25 - 0/25 = 0,60$	$\Delta P_{CB A} = 4/14 - 10/35 = 0,00$
$\Delta P_{BA C} = 21/25 - 15/25 = 0,24$	$\Delta P_{CA B} = 21/25 - 15/25 = 0,24$	$\Delta P_{CB A} = 21/36 - 15/15 = -0,42$

Abb. 2.15: Zahlenbeispiel für die Ableitung eines Kausalmodells aus Kovariationsdaten

Für eine Modellinduktion ist es als Erstes notwendig alle unbedingten und bedingten Abhängigkeiten zwischen den drei Ereignissen zu bestimmen. Die entsprechenden Ergebnisse für die Zusammenhänge zwischen den Arealen zeigt die untere Hälfte der Abbildung 2.15.

Besteht zwischen zwei Ereignissen sowohl eine unbedingte als auch eine bedingte Abhängigkeit, dann existiert ein Kausalzusammenhang zwischen diesen Ereignissen. Eine Analyse der Zusammenhänge zeigt, dass zwischen der Durchblutung in den Arealen A und B einerseits und den Arealen A und C andererseits eine bedingte und eine unbedingte Abhängigkeit besteht. Diese Areale müssen also jeweils kausal verbunden sein. Die Richtung des Zusammenhangs ist an dieser Stelle allerdings noch nicht bestimmbar.

Korrelieren zwei Ereignisse, sind aber bedingt auf das dritte Ereignis konditional unabhängig voneinander, so besteht vermutlich kein direkter Kausalzusammenhang zwischen diesen. Das dritte Ereignis ist in diesem Fall entweder eine gemeinsame Ursache der beiden oder ein zwischen diesen beiden Ereignissen liegendes vermittelndes Ereignis. Dieses spezifische Muster kommt bei den Zusammenhängen in unserem Beispiel nicht vor.

Sind zwei Ereignisse zunächst unabhängig voneinander, weisen aber bedingt auf das dritte Ereignis eine negative Abhängigkeit auf, dann sind diese beiden Ereignisse die Ursache des dritten Ereignisses. Dieses Muster trifft auf die Durchblutung in den Arealen B und C zu. Die unbedingte Kontingenz zwischen den beiden Arealen ist in beiden Richtungen Null, die bedingte Kontingenz ist jeweils negativ. Daher müssen die Aktivitäten in den Arealen B und C beide Ursachen für die Aktivität im Areal C sein. Damit ist auch die

Richtung der beiden Kausalzusammenhänge klar. Die drei Areale sind durch eine Gemeinsamer-Effekt Struktur verbunden, wobei die Areale B und C das Areal A beeinflussen.

Die Aufklärung der Struktur ist in diesem Fall nur deshalb möglich, weil die ermittelten Zusammenhänge den strukturellen Implikationen des Gemeinsamer-Effekt Modells entsprechen. Dieses weist als einziges Modell mit drei Ereignissen spezifische, von allen anderen Modellen verschiedene strukturelle Implikationen auf. In der Regel sind Kausalmodelle erst ab vier Ereignissen identifizierbar.⁴⁰ Komplexere Modelle können durch die Zerlegung in einzelne Tetraden (Teilmengen bestehend aus vier Ereignissen) und deren Analyse erschlossen werden (Glymour & Cooper, 1999).

Empirische Befunde

Bisher wurde meines Wissens noch nicht untersucht, ob Probanden in der Lage sind, allein aus dem Kovariationsmuster von Ereignissen die Struktur des dahinter liegenden Kausalmodells zu erschließen. Es erscheint aber extrem unwahrscheinlich, dass Menschen solche Aufgaben ohne die Unterstützung eines Computers bewältigen können. Wie das obige Beispiel deutlich macht, ist eine solche Aufgabe selbst bei wenigen Ereignissen sehr komplex.

Ableitung von Kausalmodellen aus Experimenten

Eine zweite mögliche Vorgehensweise bei der Aufklärung einer Kausalstruktur ist es, Experimente durchzuführen. Durch die Asymmetrie der Kausalrelation kann die Kausalstruktur aufgeklärt werden. Nur die Manipulation von Ursachen wirkt sich auch auf deren Effekte aus, eine Manipulation der Effekte hat aber keinen Einfluss auf deren Ursachen. Durch die gezielte, schrittweise Manipulation aller Ereignisse können daher die Ursachen von den Effekten unterschieden werden. Nehmen wir zum Beispiel die drei Gehirnareale X, Y und Z. Eine Möglichkeit, deren Kausalzusammenhänge aufzuklären, wäre es, die Areale unabhängig voneinander zu stimulieren und zu ermitteln, welche anderen Areale ebenfalls in der Folge aktiviert werden. Wichtig dabei ist es, dass die Stimulation per Zufall unter sonst gleichen Bedingungen durchgeführt wird. Damit wird sichergestellt, dass keine systematische Konfundierung mit anderen Faktoren gegeben ist. Angenommen, das Ergebnis dieser Experimente ist, dass eine Stimulation von X zu einer Aktivierung von Y und Z führt und eine Stimulierung von Y zu einer Aktivierung von Z. Eine Aktivierung von Z hat

⁴⁰ Eine Ermittlung des zu Grunde liegenden Kausalmodells ist auch dann nicht immer möglich. Es kann sein, dass das Modell durch die vorhandenen Kovariationen unterspezifiziert ist. (vgl. Pearl, 2000 für eine ausführliche Diskussion unter welchen Bedingungen eine valide Modellinduktion möglich ist.)

keine Aktivierung der beiden anderen Areale zur Folge. Die unten stehende Abbildung veranschaulicht die Ergebnisse der Stimulierung.

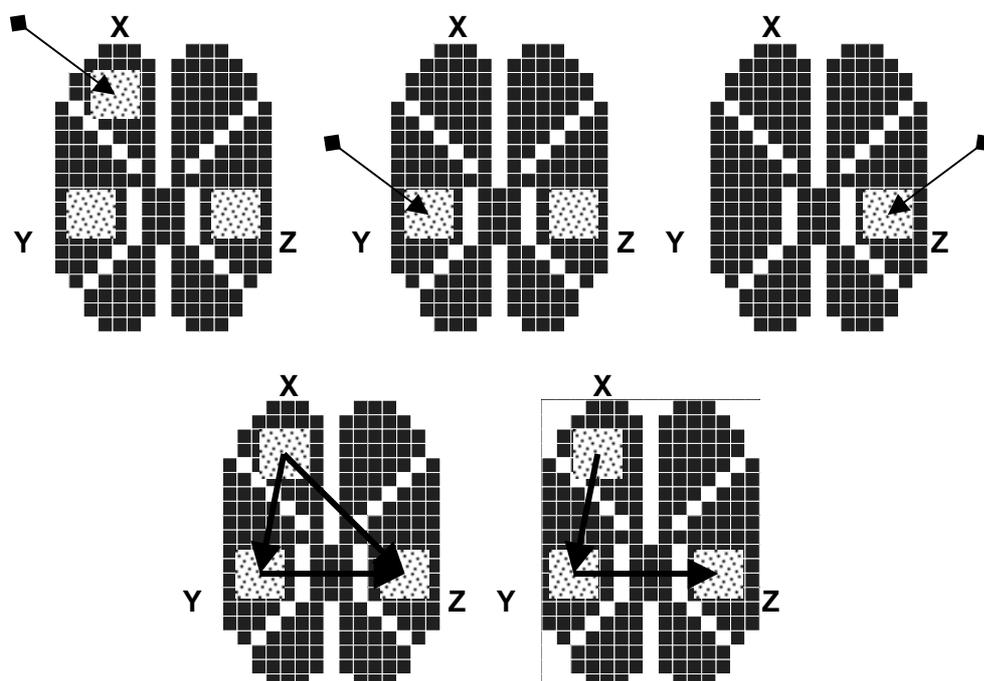


Abbildung 2.16: Aufklärung der Kausalstruktur durch Experimente; die drei oberen Abbildungen zeigen die Ergebnisse der Experimente, der Pfeil gibt das stimulierte Areal an, die gepunkteten Areale sind nach einer Stimulation aktiv; die unteren Abbildungen zeigen die damit zu vereinbarenden Kausalmodelle

Zwei Strukturen sind mit den beobachteten Veränderungen vereinbar: Entweder Areal X beeinflusst sowohl Areal Y als auch Areal Z und Areal Y beeinflusst zusätzlich Areal Z (Abbildung unten links) oder es liegt eine Kettenstruktur $X \rightarrow Y \rightarrow Z$ vor (Abbildung unten rechts). Um zwischen diesen beiden Hypothesen zu differenzieren, müssen weitere Untersuchungen durchgeführt werden. Bei diesen gilt es, einen möglichen Einfluss von Areal Y auf Areal Z auszuschließen. Eine Option hierfür wäre es, Areal Y chirurgisch zu entfernen. Führt unter diesen Bedingungen eine Stimulierung von Areal X nicht mehr zu einer Aktivierung von Areal Z, dann bilden die drei Areale eine Kettenstruktur.

Empirische Befunde

In Experimenten zum wissenschaftlichen Denken bekamen Probanden sehr häufig die Gelegenheit Experimente durchzuführen (u.a. Klahr, 2000; Kuhn et al. 1995; Schauble, 1996). Es zeigte sich, dass die Probanden diese Möglichkeit zielgerichtet nutzten und so die Zusammenhänge aufklären konnten. Allerdings waren die Aufgaben so strukturiert, dass sich nur die vorgegebenen Ursachen manipulieren ließen. Ob die Probanden auch die Effekte verändert hätten, ließ sich hiermit nicht herausfinden. Außerdem wurden stets zumindest in

Ansätzen bekannte Sachverhalte verwendet, so dass bereits aufgrund des Vorwissens Hypothesen über die Kausalstruktur ableitbar waren.

Verbindung von einzelnen Kausalzusammenhängen zu einem Kausalmodell

Ein dritter Weg zur Generierung von kausalen Strukturhypothesen ist es, mehrere einzelne Kausalzusammenhänge mit teilweise gleichen Ereignissen zu verbinden. Erfahren wir zunächst, dass Ereignis A ein Ereignis B verursacht, und lernen später, dass Ereignis A auch ein Ereignis C beeinflusst, dann kann dieses verteilte Wissen in ein einheitliches Kausalmodell integriert werden. Das Ereignis A wird zur gemeinsame Ursache der Ereignisse B und C. Häufig erwerben wir unser Wissen über Kausalzusammenhänge auf diese Weise. So lernen wir zum Beispiel schon als Kinder, dass Aspirin Kopfweh beseitigt. Später merken wir dann, dass es auch zu Magenproblemen führt. Das Wissen um diese beiden Wirkungen von Aspirin können wir nun in ein Gemeinsame-Ursache Modell integrieren.

Mit der Bildung eines Kausalmodells gelten auch dessen strukturelle Implikationen. Durch die Integration kommen also Eigenschaften hinzu, welche die beiden einzelnen Kausalzusammenhänge nicht aufweisen. Oder anders gesagt, das Kausalmodell ist mehr als die Summe seiner Kausalzusammenhänge. Diese strukturellen Implikationen wiederum erlauben es, Vorhersagen über bisher nicht beobachtete Zusammenhänge zu machen. So impliziert beispielsweise das Gemeinsame-Ursache Modell, dass die beiden Effekte kovariieren sollten. Wir werden also erwarten, dass Personen, die Aspirin eingenommen haben, zwar kein Kopfweh, dafür aber eine Magenverstimmung haben.⁴¹ Diese Erwartung ist davon unabhängig, ob wir diesen Zusammenhang bereits beobachtet haben oder nicht.

Empirische Befunde

Eine Reihe von Experimenten lieferten indirekte Evidenzen dafür, dass Probanden in der Lage sind, unterschiedliche Kausalzusammenhänge zu einem Modell zu verbinden und Schlussfolgerungen daraus abzuleiten (Waldmann, 1994). So wurde Probanden die Aufgabe gestellt, nacheinander die Auswirkung der Stärke zweier Medikamente auf die Herzrate von Patienten zu erlernen. Danach wurden sie gebeten einzuschätzen, wie sich eine Mischung der beiden Medikamente auf die Herzrate auswirkt. Das Ergebnis war, dass die Probanden für ihre Einschätzung die Stärke der Medikamente addierten, was in diesem Fall ein angemessenes Vorgehen war (Waldmann, 1994).

⁴¹ Da die Zusammenhänge nur probabilistisch sind, gilt dies natürlich nicht in jedem Einzelfall. Die meisten Personen vertragen Aspirin sehr gut und haben keine Magenbeschwerden.

Offene Fragen

In den bisher vorliegenden Experimenten wurde nicht untersucht, ob Probanden getrennt voneinander gelernte Kausalzusammenhänge in ein einheitliches Kausalmodell integrieren und daraus die strukturellen Implikationen ableiten können. Im Rahmen dieser Arbeit wurde daher eine Reihe von Experimenten mit dieser Fragestellung durchgeführt. Es bot sich an, hierzu dasselbe Versuchsdesign wie für die Experimente zur Vorhersage von nicht direkt kausal verbundenen Ereignissen zu verwenden. In beiden Fällen gilt es herauszufinden, ob Probanden sensitiv für die strukturellen Implikationen von verschiedenen Kausalmodellen sind. Bei diesem Paradigma erhielten die Probanden zunächst die Aufgabe, unabhängig voneinander zwei Kausalzusammenhänge zu erlernen. Die Kausalrelationen bildeten entweder ein Gemeinsame-Ursache Modell oder ein Gemeinsamer-Effekt Modell. Um zu überprüfen, ob sie die strukturellen Implikationen des entstandenen Modells erkennen, wurden sie dann gebeten, den Zusammenhang der beiden Ursachen bzw. den Zusammenhang der beiden Effekte einzuschätzen. Wenn die Probanden in der Lage sind, auf diesem Wege Kausalmodelle zu generieren, dann sollten sie für die unterschiedlichen strukturellen Implikationen der verschiedenen Modelle sensitiv sein. Die durchgeführten Experimente und die Ergebnisse werden in Abschnitt 3.3 ausführlich beschrieben.

2.5 Zusammenfassung

In diesem Kapitel wurde eine Vielzahl von möglichen Anwendungen von Kausalmodellen vorgestellt. Kausalmodelle müssen beim Erlernen von Kausalzusammenhängen berücksichtigt werden, sie bilden die Grundlage für Prognosen und Diagnosen und sie sind entscheidend für das Prüfen von Kausalhypothesen. In all diesen Bereichen wurden bereits psychologische Untersuchungen durchgeführt, von denen sich aber nur ein Teil direkt mit der Rolle von Kausalmodellen beschäftigt hat. Aber auch die Studien, in denen der Einfluss von Kausalmodellen nur indirekt untersucht wurde, erbrachten wichtige Erkenntnisse. Nimmt man die Befunde zusammen, dann ist meines Erachtens die generelle Behauptung gerechtfertigt, dass Personen Kausalmodelle berücksichtigen und sich den normativen Vorgaben dieser gemäß verhalten können. Eine Sensitivität für Kausalmodelle zeigte sich in allen angeführten Anwendungsfeldern. Einschränkend muss aber gesagt werden, dass dieser grundsätzlichen Kompetenz eine ganze Reihe von Performanzfaktoren entgegensteht. Je komplexer und schwieriger die Aufgabe war, desto weniger wurden Kausalmodelle berücksichtigt und desto weniger stimmten die Antworten der Probanden mit den normativ richtigen überein.

Das nun folgende Kapitel stellt meine eigenen Untersuchungen zur Frage, ob Personen sensitiv für Kausalmodelle sind, vor. Dabei wird es zunächst noch einmal darum gehen, ob Probanden bei der Bestimmung eines Kausalzusammenhangs die Struktur des umgebenden Modells berücksichtigen. Der zweite Abschnitt beleuchtet die Frage, ob Personen kausale Strukturhypothesen normativ richtig prüfen. Um dies tun zu können, müssten sie die strukturellen Implikationen der verschiedenen Modelle kennen und überprüfen. Im dritten Abschnitt wird es dann um die Bildung von Kausalmodellen aus einzelnen Zusammenhängen gehen und um die Frage, ob Probanden aus diesen Modellen Vorhersagen über die strukturell implizierten Zusammenhänge ableiten können. Damit wird gleichzeitig die Frage beantwortet, ob Probanden strukturelle Implikationen nutzen können, um nicht direkt kausal verbundene Ereignisse vorherzusagen.

3. EMPIRISCHE UNTERSUCHUNGEN

Alle in diesem Kapitel vorgestellten Experimenten beschäftigen sich mit einer übergeordneten Frage. Die Frage ist, ob Probanden für Kausalmodelle und deren Implikationen sensitiv sind. Eine Sensitivität für Kausalmodelle kann sich bei verschiedenen Aufgaben zeigen. Drei dieser Aufgaben wurden im Rahmen dieser Arbeit näher untersucht: Die Bestimmung der Stärke von Kausalzusammenhängen innerhalb eines Kausalmodells, die Prüfung von kausalen Strukturhypothesen und die Vorhersage von strukturell implizierten Zusammenhängen. Bei der Bestimmung der Stärke von Kausalzusammenhängen führt eine Sensitivität dazu, dass je nach Modell unterschiedliche statistische Indikatoren herangezogen werden. Bei der Prüfung von kausalen Strukturhypothesen zeigt sich eine Sensitivität für Kausalmodelle darin, dass sowohl die vermuteten Kausalzusammenhänge als auch die strukturellen Implikationen des Modells überprüft werden. Bei der Vorhersage strukturell implizierter Zusammenhänge führt eine Sensitivität dazu, dass je nach Modell trotz identischer Kausalzusammenhänge unterschiedliche Vorhersagen abgeleitet werden. Der Aufbau des dritten Kapitel folgt dieser Dreiteilung. Die unten stehende Abbildung zeigt die Struktur des Kapitels

3. EXPERIMENTE		
3.1 Experiment zur Bestimmung von Kausalzusammenhängen Experiment 1	3.2 Experimente zur Prüfung kausaler Strukturhypothesen Experimente 2-5 Befragungsstudie	3.3 Experimente zur Vorhersage strukturell implizierter Zusammenhänge Experimente 6-8

Abb. 3.1: Struktur Kapitel 3

Experiment 1 untersucht den Einfluss von Kausalmodellen auf die Bestimmung von Kausalzusammenhängen. Die im vorangegangenen Kapitel berichteten Befunden sprechen dafür, dass Kausalmodelle hierfür normativ richtig berücksichtigt werden. Allerdings wurde dies bisher nur für das Gemeinsame-Ursache und das Gemeinsamer-Effekt Modell gezeigt. Das Kettenmodell wurde bis dato vernachlässigt. Experiment 1 schließt diese Lücke. Das Experiment und die Befunde werden in Kapitel 3.1 geschildert.

Die Experimente 2 bis 5 gehen der Frage nach, ob Probanden in der Lage sind, kausale Strukturhypothesen zu überprüfen. Eine normativ korrekte Überprüfung setzt voraus, dass die strukturellen Implikationen der Kausalmodelle herangezogen werden. Nur wenn diese erfüllt sind, ist eine Strukturhypothese mit den Daten vereinbar. Eine Befragungsstudie klärt zusätzlich, welche Erwartungen Probanden mit verschiedenen Strukturhypothesen

verbinden. Die Experimente und die Befragungsstudie werden in Kapitel 3.2 ausführlich beschrieben.

Die Experimente 6 bis 8 untersuchen die Vorhersage strukturell implizierter Zusammenhänge. Dazu sollten Probanden verschiedene Kausalzusammenhänge zu einem einheitlichen Kausalmodell verbinden und dann den Zusammenhang zwischen den Ereignissen einschätzen, die innerhalb des Modells nicht direkt kausal verbunden sind. Eine normativ richtige Einschätzung dieser Zusammenhänge ist auf Grundlage der strukturellen Implikationen der Modelle möglich. Kapitel 3.3 stellt die Experimente und die Ergebnisse dar.

3.1 Experiment zur Bestimmung von Kausalzusammenhängen

Experiment 1

Zielsetzung dieses Experimentes ist es herauszufinden, ob Probanden bei der Bestimmung der Stärke eines Kausalzusammenhangs berücksichtigen, in welche Struktur dieser Zusammenhang eingebettet ist. Wie oben ausgeführt, sind bei unterschiedlichen kausalen Strukturmodellen verschiedene statistische Zusammenhangsmaße indikativ für die kausale Wirkung einer Ursache. In dem vorliegenden Experiment wurden alle Probanden gebeten, denselben Kausalzusammenhang zwischen zwei Ereignissen (A und B) anhand eines stets gleichen Datensatzes zu beurteilen. Zusätzlich zu diesen beiden Ereignissen lag ein drittes, kausal relevantes Ereignis (C) vor. Durch die Manipulation des kausalen Status dieses dritten Ereignisses wurden vier verschiedene Kausalmodelle erzeugt. So konnte das dritte Ereignis erstens eine Ursache von Ereignis A (Kettenmodell 1), zweitens ein weiterer Effekt der Ursache A (Gemeinsame-Ursache Modell), drittens eine weitere Ursache des Effektes B (Gemeinsamer-Effekt Modell) oder viertens ein weiterer Effekt des Ereignisses B (Kettenmodell 2) sein. Abbildung 3.2 zeigt die verschiedenen resultierenden Kausalmodelle.

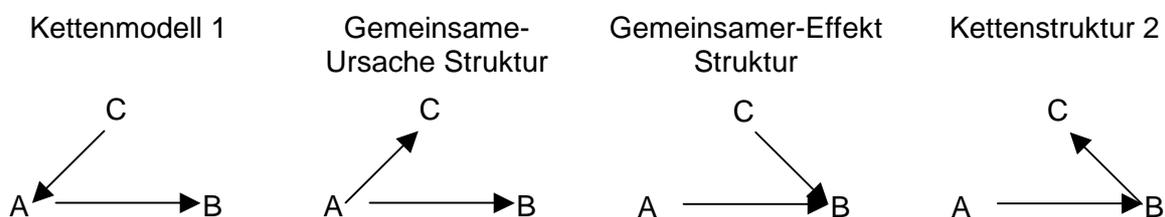


Abb. 3.2: Strukturen der in Experiment 1 verwendeten Kausalmodelle.

Welche Analyse der Daten erfordern diese verschiedenen Modelle für die Bestimmung des kausalen Einflusses von Ereignis A auf B? Als Regel gilt, dass bei der Analyse alle nicht auf die Ursache folgenden, aber den Effekt beeinflussenden Variablen

kontrolliert werden müssen, wogegen alle Ereignisse, die direkt oder indirekt von der Ursache kausal beeinflusst werden, nicht konstant gehalten werden dürfen (vgl. Eells, 1991 und Kapitel 1.3).

Bei Kettenstruktur 1 beeinflusst das Ereignis C die Ursache A. Bei Annahme der Markovbedingung ist C konditional auf A unabhängig von B. D.h., C hat keinen direkten Einfluss auf B. In diesem Fall braucht C bei der Bestimmung des Kausalzusammenhangs von A und B nicht berücksichtigt werden. Die Markovbedingung ist nicht erfüllt, wenn es möglicherweise eine Interaktion von C und A bei der Verursachung von B gibt. In diesem Fall muss C konstant gehalten werden. Ob Probanden die mögliche Interaktion von C und A in Betracht ziehen, wird das Experiment zu zeigen haben. Es ist aber eher zu erwarten, dass sie von einer Markovkette ausgehen und damit das einfachere Vorgehen ohne Kontrolle von C wählen. Sie sollten dann die unbedingte Kontingenz⁴² zwischen A und B bestimmen.

Beim Gemeinsame-Ursache Modell wird C von A direkt hervorgerufen, daher darf C nicht konstant gehalten werden. Die unbedingte Kontingenz von A auf B ist in diesem Fall indikativ für die kausale Wirkung.

Im Gegensatz hierzu erfordert das Gemeinsamer-Effekt Modell das Konstanthalten von Ereignis C. Da C eine weitere konfundierende Ursache von B ist, muss ihr Einfluss kontrolliert werden. Damit impliziert das Gemeinsamer-Effekt Modell, dass die auf C bedingte Kontingenz den tatsächlichen Kausalzusammenhang angibt.

Beim Kettenmodell 2 wird das Ereignis C von B hervorgerufen. Damit wird es indirekt von A verursacht und darf gemäß der obigen Regel nicht konstant gehalten werden. Wie beim Kettenmodell 1 (verstanden als Markovkette) und beim Gemeinsame-Ursache Modell sollte die unbedingte Kontingenz als Maß für die Stärke des kausalen Zusammenhangs verwendet werden.

Die Hypothese ist, dass die Probanden für diese strukturellen Unterschiede sensitiv sind. Sie sollten daher bei der Analyse der Daten je nach Modell unterschiedlich vorgehen und sollten so trotz identischer Datenlage und normativ richtigem Vorgehens zu anderen Einschätzungen desselben Kausalzusammenhangs kommen.

Um sicherzustellen, dass etwaige Unterschiede bei den verschiedenen Strukturen nicht auf den Inhalt des verwendeten Materials zurückzuführen sind, wurde als weiterer Faktor die Abstraktheit der verwendeten Hintergrundgeschichten variiert. Einerseits wurde ein komplett artifizierlicher chemischer Vorgang verwendet, andererseits ein plausibler Wirkmechanismus die Verdauung betreffend. Die abstrakten Inhalte bezogen sich auf die Substanz Anin (A), welche die Substanz Billon (B) hervorruft, und die Substanz Cyran (C),

⁴² Alternativ könnten die Probanden auch die kausale Power bestimmen. Beide Maße führen in diesem und den folgenden Experimenten zu denselben Vorhersagen. Aus Vereinfachungsgründen werden daher nur noch die Kontingenzen angegeben.

welche je nach Bedingung die Bildung von Anin oder Billon beeinflusst bzw. von diesen beeinflusst wird. Vorteil der abstrakten Inhalte ist es, dass für das Ereignis C stets dasselbe Ereignis (Cyran) eingesetzt werden konnte. Der Nachteil aber ist, dass mit abstrakten Ereignissen jede mögliche Struktur prima facie gleich wahrscheinlich ist. Ob Cyran Anin hervorruft oder Anin Cyran verursacht, ist aus Sicht der Probanden beides gleich plausibel. Dies könnte dazu führen, dass die Probanden die in der Instruktion angegebene Struktur nicht übernehmen. So könnte ein Teil der Probanden die Strukturangaben entweder vernachlässigen oder versuchen, aus den Daten eine Struktur neu zu entwickeln. Unter diesen Umständen wird der Einfluss des gezeigten Kausalmodells abgeschwächt werden. Bei dem Verdauungsproblem handelte es sich um die Wirkung von genetisch veränderten Früchten auf die Magen-Darm-Flora. Das Ereignis C variierte je nach Strukturmodell. So wurde als Ursache für die genetische Veränderung der Früchte Bestrahlung angegeben, die zweite Wirkung der genetischen Veränderung war eine allergische Hautrötung, die weitere Ursache für Schädigungen der Magen-Darm-Flora waren Bakterien und die Auswirkung der Schädigung der Magen-Darm-Flora waren Blähungen. Abbildung 3.3 gibt nochmals einen Überblick über alle Versuchsbedingungen. Der Vorteil der konkreten Inhalte ist, dass die Struktur unmittelbar einsichtig ist. So können Hautrötungen keine Störungen im Magen-Darmtrakt hervorrufen. Nachteil ist, dass etwaige Unterschiede zwischen den Strukturen durch das Vorwissen über mögliche Wirkmechanismen bedingt sein könnten. Es ist anzunehmen, dass bei den konkreten Inhalten der Einfluss des Kausalmodells stärker ausfällt, da die Struktur unzweifelhaft klar ist. Bei den abstrakten Inhalten könnte der Effekt wie oben erläutert etwas schwächer ausfallen. Dennoch ist zu erwarten, dass die Probanden in beiden Inhaltsbedingungen normativ richtig vorgehen.

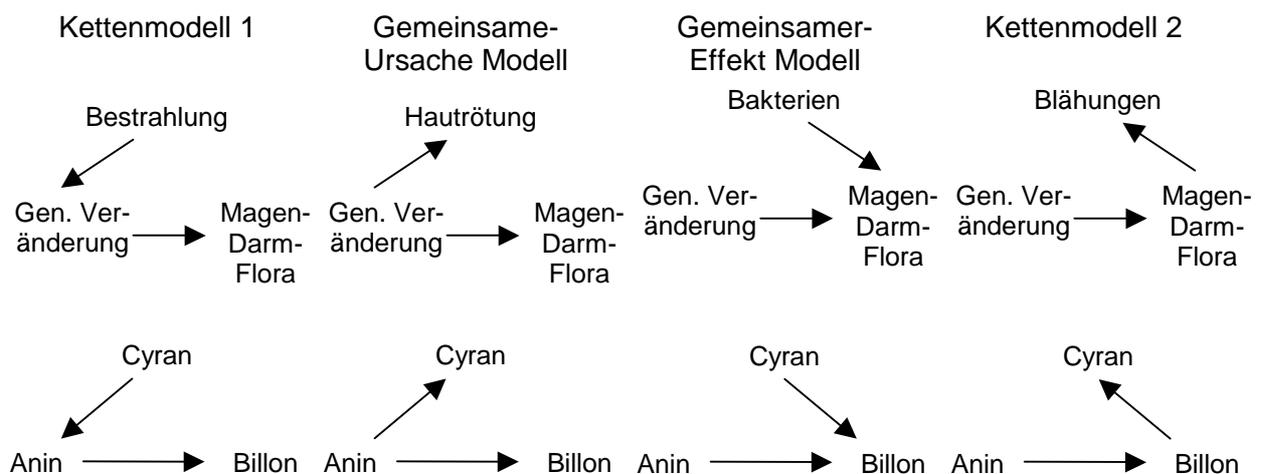


Abb. 3.3: Experimentalbedingungen Experiment 1.
Die obere Hälfte gibt die konkreten Inhalte,
die untere die abstrakten Inhalte wieder.

Probanden und Design

Sechshundneunzig Studentinnen und Studenten der Universität Göttingen nahmen an dem Experiment teil. Die Probanden verfügten zum Zeitpunkt der Untersuchung über kein einschlägiges Methodenwissen (Statistik oder Methodenlehre). Sie wurden extern angeworben und erhielten für ihre Teilnahme 10,-- DM. Unter den Probanden war eine Vielzahl von Fachbereichen vertreten, darunter nur fünf Psychologiestudierende. Diese wurden zu Beginn ihres ersten Fachsemesters getestet. Die Zuteilung in eine der acht Versuchsbedingungen erfolgte per Zufall. Die Probanden wurden einzeln untersucht.

Versuchsdurchführung und Material

Alle Teilnehmer erhielten die Aufgabe herauszufinden, ob und wie ein Ereignis A ein Ereignis B beeinflusst. Zusätzlich lag jeweils ein drittes Ereignis C vor, dessen kausale Rolle aber je nach Strukturbedingung variierte.

In der konkreten Bedingung wurden die Probanden instruiert, dass Mediziner auf der Suche nach Stoffen in der Umwelt seien, die zu Magenproblemen führen. Nun hätten die Forscher genetisch veränderte Lebensmittel in Verdacht, dass diese zu Magenproblemen führen können. Sie würden vermuten, dass die genetisch veränderten Früchte die Magen – Darm - Flora schädigen können.

Zur Manipulation des angenommenen Kausalmodells wurden die Probanden in der Instruktion an dieser Stelle über Vermutungen bezüglich eines zweiten Kausalzusammenhangs informiert. In der Bedingung mit Kettenmodell 1 wurde ihnen mitgeteilt, dass man ebenso glaube, dass eine Bestrahlung die genetische Veränderung in den Früchten verursachen könne. Zur Instruktion eines Gemeinsame-Ursache Modells wurde mitgeteilt, dass man glaube, dass die genetisch veränderten Früchte auch zu allergisch bedingten Hautrötungen führen können. Probanden in der Gemeinsamer-Effekt Bedingung lasen, dass man glaube, dass Bakterien ebenfalls die Magen – Darm – Flora schädigen können, und in der Kettenmodell 2 Bedingung, dass man glaube, dass Schädigungen der Magen – Darm – Flora zu Blähungen führen können. Zusätzlich erhielten die Probanden eine Abbildung, welche die jeweilige Kausalstruktur veranschaulichte.

Im weiteren Verlauf der Instruktion wurde mitgeteilt, dass die Forscher mit neuen Untersuchungen herausfinden wollten, ob genetisch veränderte Früchte die Magen – Darm – Flora schädigen. Dabei sei jedes Ergebnis denkbar. So könnten die Früchte die Magen – Darm – Flora mehr oder weniger stark schädigen oder aber gar keinen Einfluss haben. Die Aufgabe des Probanden sei es, mit Hilfe der Ergebnisse der Studie herauszufinden, ob und wie sich die genetisch veränderten Früchte auf die Magen – Darm – Flora auswirken.

In den abstrakten Bedingungen wurden die Probanden instruiert, dass Mediziner an Stoffwechselprozessen interessiert und deshalb auf der Suche nach Substanzen aus der Nahrung seien, die zur Bildung anderer Substanzen führen. Nun hätten die Forscher den Verdacht, dass die Substanz ANIN zur Bildung der Substanz BILLON führen könne. Sie würden vermuten, dass ANIN durch Polymerisation zur Bildung von BILLON führen könne.

Wie in der konkreten Bedingung wurde je nach Struktur ein zweiter Kausalzusammenhang erwähnt. So lasen die Probanden entweder, dass man glaube, dass CYRAN über Oxidationsprozesse zu ANIN führe (Kettenmodell 1), ANIN über Oxidationsprozesse zu CYRAN führe (Gemeinsame-Ursache Modell), CYRAN über Oxidationsprozesse zu BILLON führe (Gemeinsamer-Effekt Modell) oder BILLON über Oxidationsprozesse zu CYRAN führe (Kettenmodell 2). Auch diese Probanden erhielten eine Abbildung zur Veranschaulichung.

Im weiteren Verlauf der Instruktion erfuhren die Versuchspersonen, dass die Forscher mit neuen Untersuchungen herausfinden wollten, ob ANIN zur Bildung von BILLON führe. Dabei sei jedes Ergebnis denkbar. So könnte ANIN die Bildung von BILLON mehr oder weniger stark fördern oder aber gar keinen Einfluss haben. Die Aufgabe des Probanden sei es, mit Hilfe der Ergebnisse der Studie herauszufinden, ob und wie sich ANIN auf die Bildung von BILLON auswirke.

Nach der Instruktion wurden die Probanden gebeten, diese zusammenzufassen, um das Instruktionsverständnis zu überprüfen. Etwaige Unklarheiten wurden beseitigt. Dann erhielten alle Versuchspersonen dieselben Daten. Diese wurden in Form einer tabellarischen Liste mit 100 Fällen vorgegeben. In der ersten Spalte wurden die Beobachtungen durchnummeriert, in den weiteren Spalten war aufgeführt, ob bei der jeweiligen Beobachtung die Ereignisse A, B und C vorlagen („ja“) oder nicht („nein“). In der abstrakten Bedingung waren die Spalten mit „Cyrans“, „Anin“ und „Billon“ gekennzeichnet, in der konkreten Bedingung je nach Modell mit „Bestrahlung“, „Bakterien“, „Hautrötung“, „Blähungen“ und „genetische Veränderung“ sowie „Schädigung der Magen-Darm-Flora“. Die Probanden hatten zur Bearbeitung der Aufgabe keinerlei Zeitbegrenzung und durften sich Notizen machen. In der Regel dauerte die Bearbeitung der Aufgabe eine halbe bis dreiviertel Stunde. Zum Abschluss beantworteten alle Probanden die Frage, wie sich Anin auf die Bildung von Billon (abstrakte Bedingung) bzw. wie sich genetisch veränderte Früchte auf die Schädigung der Magen-Darm-Flora (konkrete Bedingung) auswirken. Die Probanden gaben ihr Urteil auf einer Skala von 0 (kein Einfluss) bis +10 (sehr starker Einfluss) ab.

Präsentierte Daten

Als Datensatz wurde ein sogenanntes Simpson Paradox verwendet (Cartwright, 1989; Eells, 1991; Simpson, 1951; vgl. auch Kapitel 1.3.2). Das Paradox besagt, dass ein statistischer Zusammenhang innerhalb einer Datenmenge in allen Teilmengen verschwinden oder sich sogar umkehren kann. Die im Experiment eingesetzten Daten sind ein Beispiel für das Verschwinden eines stark positiven Zusammenhangs. Abbildung 3.4 zeigt die präsentierten Daten. Die Brüche geben die relative Häufigkeit des Effektes bei An- und Abwesenheit der Ursache und des zusätzlichen Ereignisses an. So bedeutet z.B. der Eintrag 32/40, dass in 40 Fällen sowohl die Ursache als auch das Zusatzereignis vorlagen und in 32 dieser 40 Fälle auch der Effekt auftrat. In der untersten Zeile sind die resultierenden Kontingenzen aufgeführt.

	Zusatzereignis C vorhanden	Zusatzereignis C nicht vorhanden	Gesamt
Ursache vorhanden	32/40	1/10	33/50
Ursache nicht vorhanden	8/10	4/40	12/50
Kontingenz $\Delta P = P(E U) - P(E \sim U)$	$\Delta P_C = 32/40 - 8/10 = 0,00$	$\Delta P_{\sim C} = 1/10 - 4/40 = 0,00$	$\Delta P_{\text{Gesamt}} = 33/50 - 12/50 = 0,42$

Abb. 3.4: Daten zu Experiment 1. Die Zahlen geben die relative Häufigkeit des Effektes unter den jeweiligen Bedingungen an.

Unabhängig von Ereignis C weist die Ursache eine unbedingte Kontingenz von $\Delta P_{\text{Gesamt}} = +0,42$ auf. Damit scheint die Ursache eine positive Wirkung auf den Effekt zu haben. Wird aber das Ereignis C konstant gehalten, so ist die bedingte Kontingenz sowohl bei An- wie bei Abwesenheit von C gleich Null.

Wenn die Probanden wie erwartet bei einem Gemeinsamer-Effekt Modell die bedingte Kontingenz verwenden, dann sollten sie also den Zusammenhang zwischen der Ursache und dem Effekt niedrig einschätzen. Bei dem Gemeinsame-Ursache Modell und dem Kettenmodell sollten sie normativ gesehen die unbedingten Kontingenzen nehmen. Daher sollten sie denselben Zusammenhang als mittelstark einschätzen. Beim Kettenmodell 1 hängt die Einschätzung davon ab, ob die Probanden von einer Markovkette ausgehen oder nicht. Nehmen sie die Markovbedingung an, dann sollten sie die unbedingte Kontingenz verwenden und zu einen mittleren Urteil gelangen, vermuten sie dagegen eine Interaktion

der ersten Ursache mit dem vermittelnden Ereignis, dann sollten sie die bedingte Kontingenz heranziehen und eine geringe Einschätzung abgeben.

Ergebnisse

Abbildung 3.5 zeigt die Mittelwerte der Einschätzungen der Stärke des Kausalzusammenhangs für die acht Experimentalbedingungen. Für das Kettenmodell 1, das Gemeinsame-Ursache Modell und das Kettenmodell 2 ergeben sich höhere Einschätzungen als für das Gemeinsamer-Effekt Modell. Dieses Antwortmuster zeigt sich sowohl für den konkreten, wie für den abstrakten Sachverhalt.

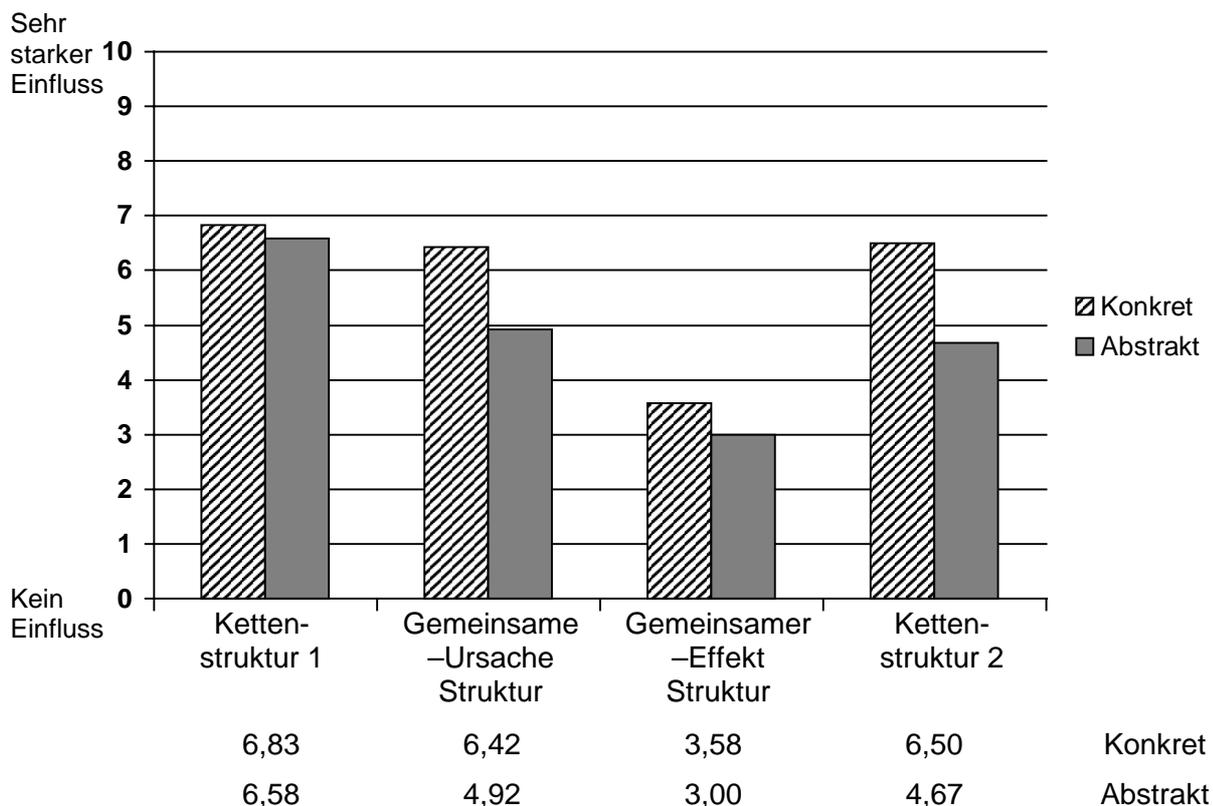


Abb. 3.5: Ergebnisse Experiment 1: Mittlere Beurteilung der Stärke des Kausalzusammenhangs.

Eine Varianzanalyse mit Struktur des Kausalmodells und Abstraktheit des Sachverhaltens als Faktoren resultierte in zwei signifikanten Haupteffekten; $F_{Struktur(3,88)} = 8,14$, $p < 0,01$, $F_{abstrakt_konkret(1,88)} = 4,25$, $p < 0,05$, $MSE = 6,12$. Die Interaktion wurde nicht signifikant ($F < 1$). Eine weitere Analyse der Daten mit dem ‚Least Significant Difference Test‘ erbrachte, dass der Effekt der Struktur darauf zurückzuführen ist, dass die Urteile in der Gemeinsamer-Effekt Bedingung signifikant niedriger sind als in den drei anderen Bedingungen. Diese wiederum unterscheiden sich nicht untereinander. Damit entsprechen die Ergebnisse vollständig den gemachten Vorhersagen.

Wie der Haupteffekt für die Struktur belegt, unterscheiden die Probanden zwischen den verschiedenen Kausalmodellen. Die niedrigen Urteile in der Gemeinsamer-Effekt Bedingung sprechen dafür, dass die Probanden das zusätzliche Ereignis konstant gehalten haben und die bedingte Kontingenz als Indikator für die Stärke des Kausalzusammenhangs herangezogen haben.⁴³ Die Urteile in den anderen drei Bedingungen sind signifikant höher und liegen im mittleren Bereich der Ratingskala. Eine solche Einschätzung sollte resultieren, wenn das zusätzliche Ereignis nicht konstant gehalten wird und die unbedingte Kontingenz als Maß verwendet wird. Damit entspricht das Antwortmuster den normativen Implikationen der Kausalmodelle. Beim Gemeinsame-Ursache Modell und dem Kettenmodell 2 gibt die unbedingte Kontingenz die Stärke des Kausalzusammenhangs an, im Gegensatz dazu ist es beim Gemeinsamer-Effekt Modell die bedingte Kontingenz. Beim Kettenmodell 1 ist die Verwendung der unbedingten Kontingenz nur legitim, wenn angenommen wird, dass die erste Ursache lediglich über das mittlere Ereignis auf den Effekt wirken kann, d.h. wenn eine Markovkette vorliegt. Das Urteil der Probanden spricht eindeutig für ein solches Verständnis der kausalen Kette. Das normative Vorgehen in allen vier Bedingungen war zudem nicht davon abhängig, ob der Sachverhalt konkret oder abstrakt war. Es zeigte sich keine Interaktion.

Doch auch für die Abstraktheit ergab sich ein signifikanter Effekt. Der konkrete Sachverhalt mit seinen plausiblen Kausalstrukturen führte zu deutlich höheren Urteilen in den Bedingungen, in denen die hohe unbedingte Kontingenz indikativ für den Kausalzusammenhang war.

Diskussion

Die Fragestellung war, ob Probanden bei der Beurteilung der Stärke eines Kausalzusammenhangs das Kausalmodell berücksichtigen, in das der Zusammenhang eingebettet ist. Dazu wurden sie gebeten, den Kausalzusammenhang zwischen zwei Ereignissen einzuschätzen, wobei zusätzlich jeweils ein weiteres Ereignis gegeben war, dessen kausale Rolle variiert wurde. Für die vier resultierenden Kausalmodelle ergaben sich folgende Ergebnisse. Wenn dieses dritte Ereignis eine weitere Ursache des Zieleffektes war, dann kontrollierten die Probanden dieses Ereignis und verwendeten die bedingte Kontingenz als Indikator für Kausalität. War das dritte Ereignis eine weitere Ursache der Zielursache, ein weiterer Effekt der Zielursache oder ein Effekt des Zieleffektes, dann vernachlässigten die Probanden dieses Ereignis und nahmen die unbedingte Kontingenz als Maß. Damit verhielten sich die Teilnehmenden normativ.

⁴³ Dass die Urteile dennoch von Null abweichen, ist ein häufig beobachteter Befund (siehe Buehner & Cheng, 1997; Shanks, 1993). Probanden scheinen allgemein Schwierigkeiten damit zu haben, eine Nullkontingenz zu erkennen, wenn der Effekt auch in Abwesenheit der Ursache auftritt (vgl. auch Alloy & Tabachnik, 1984)

Bisherige Forschungsarbeiten hatten sich v.a. auf das Gemeinsamer-Effekt Modell und die Frage, ob Probanden den Einfluss weiterer Ursachen kontrollieren, konzentriert (Baker et al., 1993; Spellman, 1996a,b; Waldmann & Hagmayer, 1995; siehe Kapitel 2.1 für einen Überblick). Die vorliegenden Befunde stimmen mit den Ergebnissen dieser Arbeiten überein. Der Einfluss konfundierender Ursachen wird kontrolliert. Die bedingte Kontingenz in Abwesenheit anderer Ursachen wird als Maß für die kausale Wirksamkeit herangezogen.

Eine Reihe anderer Befunde deuten aber darauf hin, dass die Sensitivität für Kausalmodelle von Anforderungen an die kognitive Verarbeitungskapazität abhängig ist. Sind die Daten komplex, die Präsentation unübersichtlich oder die Kausalstruktur nicht eindeutig klar, wird der Einfluss zusätzlicher Ursachen nicht beachtet (Fiedler et al.; 2000, Reips, 1997; Waldmann, 2000; Waldmann & Hagmayer, subm.). In diesem Experiment wurden die Anforderungen an die Verarbeitungskapazität gering gehalten. So lagen die Daten den Probanden durchgehend als Liste vor und sie durften sich Notizen machen. Auch die Bearbeitungszeit war nicht eingeschränkt. Ob sich eine Sensitivität für Kausalmodelle auch bei höheren kognitiven Anforderungen zeigt, muss weitere Forschung klären.

Der signifikante Inhaltseffekt passt ebenfalls gut zu den bisher vorliegenden Befunden. So konnte gezeigt werden, dass Probanden Kausalzusammenhänge bei gleicher Kontingenz stärker einschätzen, wenn ihnen ein plausibler Mechanismus bekannt ist, der die Ereignisse verbindet (Koslowski, 1996). Die fehlende Interaktion belegt aber, dass der Einfluss des Kausalmodells von diesem Vorwissenseffekt unabhängig ist.

Was bedeutet das gefundene Ergebnis für unsere Frage nach der Sensitivität für kausale Strukturmodelle?

Als erstes zeigt sich ein beeindruckend normatives Vorgehen der Probanden. Verschiedene Kausalmodelle werden richtig unterschieden und es werden die für die Struktur korrekten statistischen Indikatoren für den Kausalzusammenhang ausgewählt. Das bedeutet, dass die Personen die Kausalmodelle richtig anwenden können.

Als zweites zeigt sich aber geringes Maß an Sensitivität für strukturelle Implikationen. In den Daten liegt ein Zusammenhang zwischen der Ursache A und dem Effekt B vor, welcher aber durch das Konstanthalten des Zusatzereignisses C verschwindet. Dieses Verschwinden spricht eindeutig dafür, dass kein direkter kausaler Zusammenhang zwischen der Ursache und dem Effekt vorliegt. Die im Experiment präsentierten Daten widersprechen also den strukturellen Implikationen aller vier vorgegebener Kausalmodelle. Dies scheint aber keiner der Versuchspersonen aufgefallen zu sein. In der an das Experiment anschließenden Befragung haben lediglich zwei der 96 Probanden Zweifel an dem in der Instruktion eingeführten Kausalmodell geäußert. Keiner der beiden war aber auf Nachfrage in der Lage, ein richtiges Modell zu nennen. Ganz ähnliche Befunde berichten auch eine

Reihe anderer Forscher (Fiedler et al., in Vorbereitung; Schaller, 1992; Waldmann & Hagmayer, 1995), welche ebenfalls das Simpson Paradox für die Konstruktion von Daten eingesetzt haben. In keinem der Experimente sei den Probanden aufgefallen, dass die Daten die Implikationen der vorgegebenen Struktur nicht erfüllt hätten.

Welche Modelle sind mit den gezeigten Daten vereinbar? Drei Modelle implizieren eine unbedingte Abhängigkeit zwischen A und B, welche durch Konstanthalten von C verschwindet: Ein Gemeinsame-Ursache Modell mit C als Ursache oder Kettenmodell mit C als intermittierendem Ereignis. Abbildung 3.6 zeigt die mit den präsentierten Daten vereinbaren Modelle.

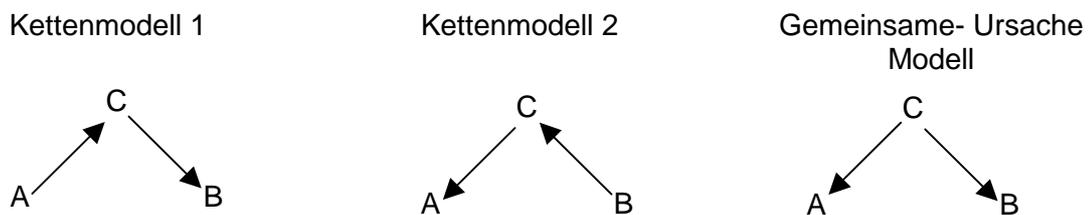


Abb. 3.6: Strukturen, welche die in Experiment 1 vorliegenden Abhängigkeiten implizieren

Die Stärke der beiden Kausalzusammenhänge kann bei jedem dieser drei Modelle über die unbedingten Kontingenzen bestimmt werden. Sie würden in diesem Fall $\Delta P_{AC} = \Delta P_{CA} = 0,60$ und $\Delta P_{CB} = \Delta P_{BC} = 0,70$ betragen. Jedes der Modelle sagt bei diesen Kontingenzen einen Zusammenhang von $\Delta P = 0,42$ zwischen A und B vorher. Dies ist der Zusammenhang der tatsächlich zwischen den Ereignissen bestand.

Eine mögliche Erklärung dafür, dass die Verletzung der strukturellen Implikationen nicht erkannt wurde, mag darin liegen, dass die Prüfung des Modells nicht Bestandteil der Aufgabe war. Es erscheint daher aus Sicht der Versuchsperson rational, das angegebene Modell zu übernehmen und sich auf die Ermittlung der Stärke des Zusammenhangs zu konzentrieren. Die Experimente im folgenden Abschnitt werden sich mit der Prüfung von kausalen Strukturhypothesen beschäftigen. Dann wird es sich zeigen, ob Probanden sensitiv für strukturelle Implikationen von Kausalmodellen sind.

3.2 Experimente zur Prüfung kausaler Strukturhypothesen

In Experiment 1 wurden Probanden gebeten zu ermitteln, ob zwischen zwei Ereignissen ein Kausalzusammenhang besteht oder nicht. Die Aufgabe, eine einfache Kausalhypothese zu überprüfen, verlangt ebenfalls eine Bestimmung des Kausalzusammenhangs (vgl. Kapitel 2.3). Die Hypothese wird dann angenommen, wenn der vermutete Zusammenhang vorliegt und abgelehnt, wenn er nicht da ist. Die umgebende Kausalstruktur gibt dabei vor, welcher statistische Zusammenhang als Entscheidungskriterium heranzuziehen ist.

Von diesen einfachen Hypothesen unterscheiden sich kausale Strukturhypothesen ganz wesentlich. Sie fragen nicht nach dem Vorliegen einzelner Zusammenhänge, sondern nach der Struktur. Wie in Kapitel 2.3.3 dargelegt, ist dabei die Struktur mehr als die Summe der Einzelzusammenhänge. Die Strukturen unterscheiden sich in ihren Implikationen bezüglich der Zusammenhänge zwischen nicht direkt kausal verbundenen Ereignissen. Eine Strukturhypothese sollte nur angenommen werden, wenn alle vermuteten Kausalzusammenhänge durch statistische Zusammenhänge bestätigt werden *und* zusätzlich die strukturellen Implikationen erfüllt sind.

Ziel der folgenden Experimente war es zu untersuchen, ob Probanden bei der Prüfung von Strukturhypothesen die unterschiedlichen Implikationen verschiedener Strukturen als Kriterium nutzen. In vier Experimenten wurden Probanden gebeten, eine oder mehrere verschiedene Hypothesen anhand vorliegender Daten zu beurteilen. Von der Vielzahl verschiedener Strukturmodelle wurden das Gemeinsame-Ursache und das Gemeinsamer-Effekt Modell als die beiden wichtigsten Vertreter ausgewählt. Abbildung 3.7 zeigt erneut die beiden Strukturmodelle sowie ihre strukturellen Implikationen.

Die Gemeinsame-Ursache Struktur impliziert eine Abhängigkeit zwischen den beiden von der einen Ursache A hervorgerufenen Effekten B und C. Da diese Abhängigkeit aber vollständig auf A zurückzuführen ist, sollten B und C durch Konstanthalten von A unabhängig werden. Im Gegensatz hierzu impliziert die Gemeinsamer-Effekt Struktur keine Abhängigkeit zwischen den Ursachen B und C. Wird aber der gemeinsame Effekt A konstant gehalten, so sollte dies zu einer Abhängigkeit von B und C führen. Die beiden Strukturen sind also anhand dieser Implikationen eindeutig voneinander zu unterscheiden.

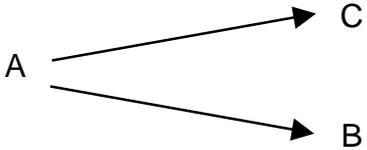
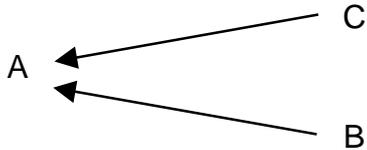
	Gemeinsame-Ursache Modell	Gemeinsamer-Effekt Modell
		
Strukturelle Implikationen	Abhängigkeit von B und C $P(C B) \neq P(C \sim B)$ Konditional auf A: Unabhängigkeit von B und C $P(C A.B) = P(C A)$	Unabhängigkeit von B und C $P(C B) = P(C \sim B)$ Konditional auf A: Abhängigkeit von B und C $P(C A.B) \neq P(C A)$
Bayesianisches Modell	$P(A,B,C) = P(B A) \cdot P(C A) \cdot P(A)$	$P(A,B,C) = P(A B,C) \cdot P(B,C)$ Angenommen, dass A und B unabhängig sind und C unabhängig voneinander beeinflussen: $P(A,B,C) = [1-(1-P(A B)) \cdot (1-P(A C))]$ $\cdot P(B) \cdot P(C)$

Abb. 3.7: Strukturelle Implikationen und Bayesianische Modelle für die in den Experimenten 2 bis 5 untersuchten Strukturmodelle.

Die präsentierten Daten waren so konstruiert, dass sie die Hypothesen entweder vollständig bestätigten, ansatzweise bestätigten oder ganz widerlegten. Eine vollständige Bestätigung war dann gegeben, wenn sowohl alle vermuteten Kausalzusammenhänge vorlagen als auch die strukturellen Implikationen erfüllt waren. Als ansatzweise bestätigt konnte eine Hypothese gelten, wenn zumindest die angenommenen Zusammenhänge gegeben waren. Das Nichtzutreffen der strukturellen Implikationen deutete in diesem Fall allerdings darauf hin, dass entweder das Modell unvollständig oder aber die angegebene Richtung der Kausalpfeile falsch war. Eine Hypothese war eindeutig widerlegt, wenn nicht einmal die vermuteten Kausalzusammenhänge vorlagen.

Wie sollte eine normativ korrekte und damit valide Prüfung einer Strukturhypothese ablaufen? Für eine Prüfung der strukturellen Implikationen sollten als Erstes unbedingte und bedingte Abhängigkeiten zwischen allen Ereignissen bestimmt werden. Die Ereignisse, welche über Kausalpfeile verbunden sind, sollten sich dabei als abhängig erweisen. Nicht direkt kausal verknüpfte Ereignisse sollten gemäß den strukturellen Implikationen des jeweiligen Modells voneinander abhängig bzw. unabhängig sein. Wenn die Implikationen nicht erfüllt sind, dann ist das Modell als falsch oder unvollständig abzulehnen.

Die ebenfalls in Abbildung 3.7 angeführten Bayesianischen Modelle liefern einen weiteren Ansatz zur Überprüfung von Strukturhypothesen. Dabei setzt die Formalisierung der beiden Modelle das Zutreffen der strukturellen Implikationen voraus. Wie erfolgt eine Strukturprüfung mit Hilfe der Bayesianischen Modelle? Zunächst werden die innerhalb des Modells verwendeten Wahrscheinlichkeiten empirisch bestimmt. Das heißt, es wird sowohl die gesamte Wahrscheinlichkeitsverteilung ($P(A,B,C)$) ermittelt als auch die bedingten Wahrscheinlichkeiten der Effekte und Basisraten für die Ursacheereignisse. Danach wird überprüft, ob sich anhand der Modellgleichung über die bedingten Wahrscheinlichkeiten und Basisraten die gesamte Wahrscheinlichkeitsverteilung reproduzieren lässt. Ist dies der Fall, trifft das Modell zu.

Die folgenden vier Experimente untersuchen, ob Probanden die strukturellen Implikationen beim Prüfen kausaler Strukturhypothesen als Kriterium verwenden und so zu normativ richtigen Urteilen kommen.

3.2.1 Experiment 2

Mit diesem Experiment wurde untersucht, ob Probanden über ein zumindest grobes Verständnis struktureller Implikationen verfügen und dies bei der Überprüfung verschiedener Strukturhypothesen einsetzen. Unter grobem Verständnis ist dabei zu verstehen, dass sie zumindest hinsichtlich der unbedingten Abhängigkeiten einen Unterschied zwischen verschiedenen Modellen machen. Die Probanden sollten nach dieser Annahme bei einem Gemeinsame-Ursache Modell erwarten, dass die Effekte kovariieren und bei einem Gemeinsamer-Effekt Modell erwarten, dass die Ursachen unabhängig sind.

Zur Untersuchung dieser Frage wurden den Probanden sukzessive vier verschiedene Strukturhypothesen vorgelegt, welche sie nacheinander anhand eines Datensatzes zu überprüfen hatten. Die Strukturen bestanden aus drei fiktiven chemischen Ereignissen, der Halogenisierung von Carboxiden (Halo.), der Kristallisation von Hygrosulfiden (Krist.) und der Ionisierung von Argobenziden (Ion.). Abbildung 3.8 stellt die vier Strukturhypothesen dar.

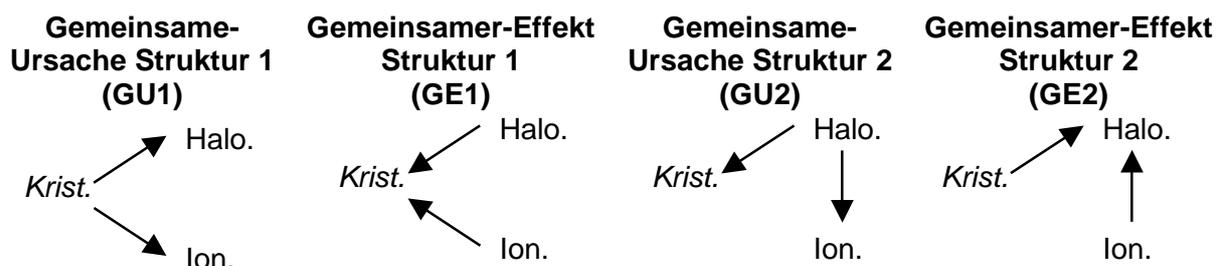
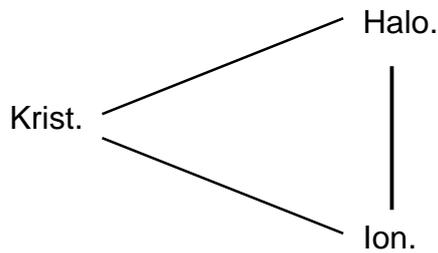
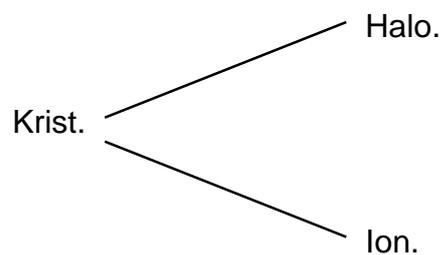


Abb 3.8: Strukturhypothesen für Experiment 2.
 Krist. = Kristallisation, Halo. = Halogenisierung, Ion. = Ionisierung

Die vier Hypothesen unterschieden sich in zwei Gesichtspunkten. Je zwei waren Gemeinsame-Ursache und Gemeinsamer-Effekt Modelle. Zusätzlich wurde variiert, welches der drei Ereignisse die gemeinsame Ursache bzw. der gemeinsame Effekt war. Während es bei den Hypothesen mit der Nummer eins die Kristallisation war, hatte bei den Strukturen mit der Ziffer zwei die Halogenisierung diese zentrale Stellung inne.

Neben den Strukturhypothesen wurden die gezeigten Daten manipuliert. Diese wiesen entweder statistische Zusammenhänge zwischen allen drei Ereignissen auf (Pro Gemeinsame-Ursache) oder nur zwischen der Kristallisation und der Halogenisierung sowie der Kristallisation und der Ionisierung (Pro Gemeinsamer Effekt 1). In letzterem Fall waren die Halogenisierung und die Ionisierung unabhängig voneinander. Abbildung 3.9 stellt die Zusammenhänge qualitativ dar.

Daten pro Gemeinsame-Ursache**Daten pro Gemeinsamer-Effekt 1**

*Abb. 3.9: Datensätze zu Experiment 2.
Durchgezogene Linien geben statistische Abhängigkeiten an*

Welche Urteile werden für die einzelnen Strukturhypothesen erwartet? Gehen wir zunächst davon aus, dass die Probanden wie vermutet ein grobes Verständnis der strukturellen Implikationen haben und auf dieser Grundlage entscheiden. Dann sollten sie bei den beiden Gemeinsame-Ursache Hypothesen prüfen, ob die Ursache mit den beiden Effekten und die Effekte untereinander zusammenhängen. Dies ist bei den Daten ‚Pro Gemeinsame-Ursache‘ der Fall. Die beiden Gemeinsame-Ursache Hypothesen sollten also bei diesen Daten angenommen werden.

In den Daten ‚Pro Gemeinsamer-Effekt 1‘ sind die Halogenisierung und die Ionisierung unabhängig voneinander. Das bedeutet, dass für die Gemeinsame-Ursache Hypothese 1 die strukturellen Implikationen nicht erfüllt sind. Diese Hypothese würde eine Kovariation der beiden Ereignisse implizieren. Bei der Gemeinsame-Ursache Hypothese 2 wird ein Kausalzusammenhang zwischen der Halogenisierung und der Ionisierung erwartet. Dieser fehlt bei den Daten ‚Pro Gemeinsamer Effekt 1‘ vollständig. Beide Gemeinsame-Ursache Hypothesen sollten also bei den Daten ‚Pro Gemeinsamer-Effekt 1‘ abgelehnt werden.

Im Gegensatz hierzu sollten die Probanden bei der Prüfung der Gemeinsamer-Effekt Strukturhypothesen darauf achten, ob die beiden vermuteten Kausalzusammenhänge vorliegen und die beiden Ursachen unabhängig voneinander sind. Bei den Daten ‚Pro Gemeinsamer-Effekt 1‘ sind die Halogenisierung und die Ionisierung unabhängig voneinander, weisen aber beide einen Zusammenhang mit der Kristallisation auf. Die Probanden sollten daher die gleichnamige Hypothese annehmen. Die Gemeinsamer-Effekt Hypothese 2 sollten sie aber ablehnen, da kein Zusammenhang zwischen der Ionisierung und der Halogenisierung gegeben ist, obwohl laut Hypothese eine Relation vorliegen sollte. Zusätzlich korrelieren die beiden vermeintlichen Ursachen Ionisierung und Kristallisation stark miteinander, was ebenfalls der Hypothese widerspricht.

Bei den Daten ‚Pro Gemeinsame-Ursache‘ kovariieren alle Ereignisse miteinander. Dies widerspricht den strukturellen Implikationen von Gemeinsamer-Effekt Modellen. Diese beiden Hypothesen sollten also bei den Daten ‚Pro Gemeinsame-Ursache‘ zurückgewiesen werden.

Was ist zu erwarten, wenn die Probanden kein Verständnis für strukturelle Implikationen mitbringen? In diesem Fall ist davon auszugehen, dass die Probanden lediglich die vermuteten Kausalzusammenhänge auf ihr Vorliegen überprüfen. Bei den Daten ‚Pro Gemeinsame-Ursache‘ liegen bei allen Modellen die angenommenen Zusammenhänge vor. Die Probanden sollten daher bei diesen Daten alle Strukturhypothesen annehmen. Bei den Daten ‚Pro Gemeinsamer-Effekt 1‘ sollten sie die Gemeinsame-Ursache- und Gemeinsamer-Effekt Strukturhypothese 1 annehmen, da beide von Kausalverbindungen zwischen denselben Ereignissen ausgehen. Diese Vermutungen werden durch entsprechende Kovariationen in den Daten bestätigt. Die Strukturhypothesen mit der Kennziffer zwei sollten sie aber ablehnen, da jeweils einer der beiden vermuteten Kausalzusammenhänge nicht vorhanden ist. Abbildung 3.10 gibt nochmals einen Überblick über die verschiedenen Antwortmuster.

	Prüfen struktureller Implikationen				Prüfen der Zusammenhänge			
	GU1 $I \leftarrow K \rightarrow H$	GE1 $I \rightarrow K \leftarrow H$	GU2 $I \leftarrow H \rightarrow K$	GE2 $I \rightarrow H \leftarrow K$	GU1 $I \leftarrow K \rightarrow H$	GE1 $I \rightarrow K \leftarrow H$	GU2 $I \leftarrow H \rightarrow K$	GE2 $I \rightarrow H \leftarrow K$
Daten pro Gemeinsamer- Effekt 1	↓	↑	↓	↓	↑	↑	↓	↓
Daten pro Gemeinsame- Ursache	↑	↓	↑	↓	↑	↑	↑	↑

Abbildung 3.10: Erwartete Antwortmuster für verschiedene Prüfstrategien.
 GU = Gemeinsame-Ursache Hypothese, GE = Gemeinsamer-Effekt Hypothese;
 ↑ = Zustimmung, ↓ = Ablehnung der jeweiligen Hypothese.

Probanden und Design

An dem Experiment nahmen 48 Studierende der Universität Tübingen teil. Die Probanden waren zur Hälfte Studierende verschiedener Fachbereiche und erhielten für ihre Teilnahme DM 10,-, die andere Hälfte waren Psychologiestudenten verschiedener Fachsemester und erhielten eine Versuchspersonenstunde testiert. Die Probanden wurden per Zufall einer der beiden Datenbedingungen zugeteilt. Alle Probanden beurteilten nacheinander die vier Strukturhypothesen, wobei die Abfolge mittels eines lateinischen Quadrates ausbalanciert wurde. Das Experiment wurde im Einzelversuch durchgeführt.

Versuchsablauf und Material

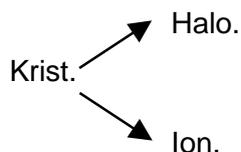
Die Probanden wurden instruiert, dass Chemiker dabei seien, die chemischen Prozesse der oberen Atmosphärenschichten zu untersuchen. Nun glaubten sie vor einem Durchbruch zu stehen. Sie seien auf drei Ereignisse gestoßen, von denen sie annähmen, dass sie einen zentralen Reaktionsmechanismus bilden. Die Ereignisse seien die Halogenisierung von Carboxiden, die Kristallisation von Hygrosulfiden und die Ionisierung von Argobenziden. Die Forscher wüssten derzeit noch nichts über die Kausalstruktur, welche die drei Ereignisse miteinander verbindet. Sie hätten aber eine Hypothese über die zu Grunde liegende Kausalstruktur, d.h. darüber, welche Ursache-Wirkungsbeziehungen zwischen den drei Ereignissen beständen. Danach wurde die erste Strukturhypothese den Probanden auf einer separaten Abbildung veranschaulicht. Die Probanden wurden darauf hingewiesen, dass die Forscher nur die gezeigten Ursache-Wirkungsbeziehungen annehmen würden. Die Aufgabe der Probanden sei es zu überprüfen, ob die Hypothese der Forscher zuträfe. Sie sollten anhand vorliegender Messdaten ermitteln, ob die Ursache-Wirkungsbeziehungen so seien, wie es die Forscher vermuteten.

Nach der Instruktion wurde deren Verständnis überprüft. Daraufhin wurde den Versuchsteilnehmern eine zweiseitige tabellarische Liste mit Daten ausgehändigt. In der ersten der vier Spalten wurden die einzelnen Erhebungen durchnummeriert, in den folgenden Spalten war eingetragen, welche der drei Ereignisse bei jeder Messung gefunden wurden. Die Liste umfasste insgesamt 96 Einzelmessungen. Für die Bearbeitung der Daten hatten die Probanden keine Zeitbeschränkung, Notizen waren ausdrücklich erlaubt. Nach Analyse der Daten wurden die Probanden gefragt, ob die Ursache-Wirkungszusammenhänge so seien, wie sie die Abbildung zeigt. Sie gaben ihre Antwort durch Ankreuzen auf einer offenen Skala mit den Endpunkten ‚sicher nicht‘ und ‚sicher ja‘. Zu ihrer Antwort erhielten die Probanden kein Feedback. Dann wurde ihnen die Hypothese einer anderen Forschergruppe vorgelegt und sie sollten diese auf Grundlage derselben Daten beurteilen. Wiederum gab es keinerlei Rückmeldung zu der Antwort. So beurteilten die Probanden sukzessive die vier Strukturhypothesen.

Präsentierte Daten

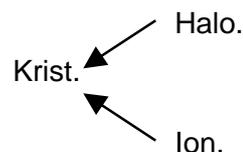
Es wurden zwei Datensätze konstruiert. Der erste basierte auf der Gemeinsame-Ursache Struktur 1, der andere auf der Gemeinsamer-Effekt Struktur 1. Abbildung 3.11 gibt die genauen Zahlen wieder.

Daten pro Gemeinsame-Ursache



	Ion.		~Ion.	
	Halo.	~Halo.	Halo.	~Halo.
Krist.	30	6	6	6
~Krist.	6	6	6	30

Daten pro Gemeinsamer-Effekt 1



	Ion.		~Ion.	
	Halo.	~Halo.	Halo.	~Halo.
Krist.	24	12	12	0
~Krist.	0	12	12	24

*Abb. 3.11: Daten zu Experiment 2.
Die Zahlen geben die Häufigkeit der jeweiligen Ereigniskombination an.*

Die Konstruktion erfolgte nach folgenden Prinzipien. Zunächst wurde die Stärke der Kausalbeziehungen parallelisiert. Dabei wurden die Kontingenzen gleichgesetzt, welche bei dem jeweiligen Modell indikativ für die Stärke des Kausalzusammenhangs sind, d.h. beim Gemeinsame-Ursache Modell betrug die unbedingte Kontingenz $\Delta P = 0,5$ und Gemeinsamer-Effekt Modell war die bedingte Kontingenz $\Delta P_C = 0,5$.

Des Weiteren waren die beiden Ursachen bei der Gemeinsamer-Effekt Struktur unabhängig voneinander, während im Gegensatz dazu die beiden Effekte bei der Gemeinsame-Ursache Struktur miteinander kovariierten. Um diese Abhängigkeit salienter für die Probanden zu machen, wurde dieser Zusammenhang gegenüber dem normativ implizierten leicht erhöht. Bei einer Kontingenz von $\Delta P = 0,5$ wäre eine Korrelation der Effekte von $\Phi = \Delta P = 0,25$ zu erwarten, diese wurde auf $\Phi = 0,5$ erhöht. Hierfür gab es zwei Gründe. Erstens könnte eine so niedrige Korrelation wie $\Phi = 0,25$ übersehen werden, wodurch die beiden Datensätzen äquivalent würden. Zweitens sind durch die Erhöhung die Beziehungen zwischen allen drei Ereignissen gleich. Dadurch sind bei den Daten ‚Pro Gemeinsame-Ursache‘ alle vermuteten Zusammenhänge bei allen vier Strukturhypothesen mit gleicher Stärke vorhanden. Prüfen die Probanden also nur das Vorhandensein der vermuteten Zusammenhänge, so sollten sie bei diesen Daten, wie oben bereits ausgeführt, allen Hypothesen gleichermaßen zustimmen.

Schließlich wurde die Verteilung der Häufigkeiten symmetrisch angelegt. Dies hat den Vorteil, dass Kontingenzen und Korrelationen einander entsprechen. Das heißt, wenn die Kontingenz von Kristallisation auf Halogenisierung $\Delta P = 0,5$ ist, dann ist auch die Φ -Korrelation von beiden 0,5.

Die beiden zuletzt genannten Konstruktionsprinzipien haben neben den genannten Vorteilen auch einen Nachteil. Durch sie entfernen sich die Datensätze von den normativ zu erwartenden. Weder sollte eine Symmetrie noch ein so hoher Zusammenhang zwischen den Effekten bestehen. Dennoch überwiegen meines Erachtens die Vorteile. Selbst wenn die Probanden vollständig normativ vorgehen sollten, so verändern sich die Vorhersagen für die Prüfung anhand struktureller Implikationen nicht wesentlich. Für die Daten ‚Pro Gemeinsamer-Effekt 1‘ ändert sich nichts. Lediglich die Zustimmung zu den beiden Gemeinsame-Ursache Hypothesen bei den Daten ‚Pro Gemeinsame-Ursache‘ könnte wegen der überhöhten Korrelation eventuell leicht zurückgehen. Die beiden Hypothesen erklären nichtsdestotrotz diese Daten immer noch besser als die Gemeinsamer-Effekt Hypothesen. Abbildung 3.12 fasst nochmals alle Zusammenhänge zwischen den Ereignissen zusammen.

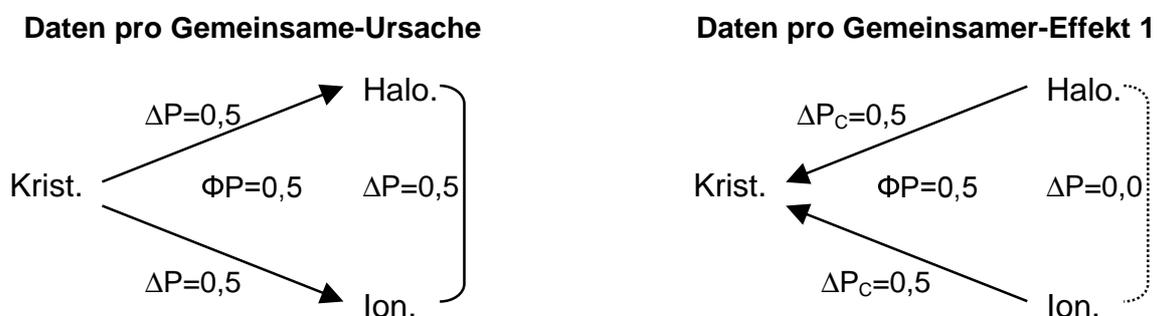


Abb. 3.12: Zusammenhänge zwischen den Ereignissen in den Daten zu Experiment 2; Krist. = Kristallisation, Halo. = Halogenisierung, Ion. = Ionisierung

Ergebnisse

Die Antworten auf der offenen Skala wurden durch Ausmessen in numerische Werte von 0 (sicher nicht) bis 100 (sicher ja) übertragen. Abbildung 3.13 zeigt die Mittelwerte für die Antworten in den jeweiligen Bedingungen.

Wenn die Daten für die Gemeinsamer-Effekt Struktur 1 sprachen (schraffierte Balken), dann fanden die Gemeinsame-Ursache Strukturhypothese 1 und die Gemeinsamer-Effekt Strukturhypothese 1 die Zustimmung der Probanden. Die beiden anderen Hypothesen wurden zurückgewiesen. Wenn dagegen die Daten für eine Gemeinsame-Ursache Struktur sprachen (ausgefüllte Balken), wurden alle vier Strukturhypothesen als zutreffend angesehen.

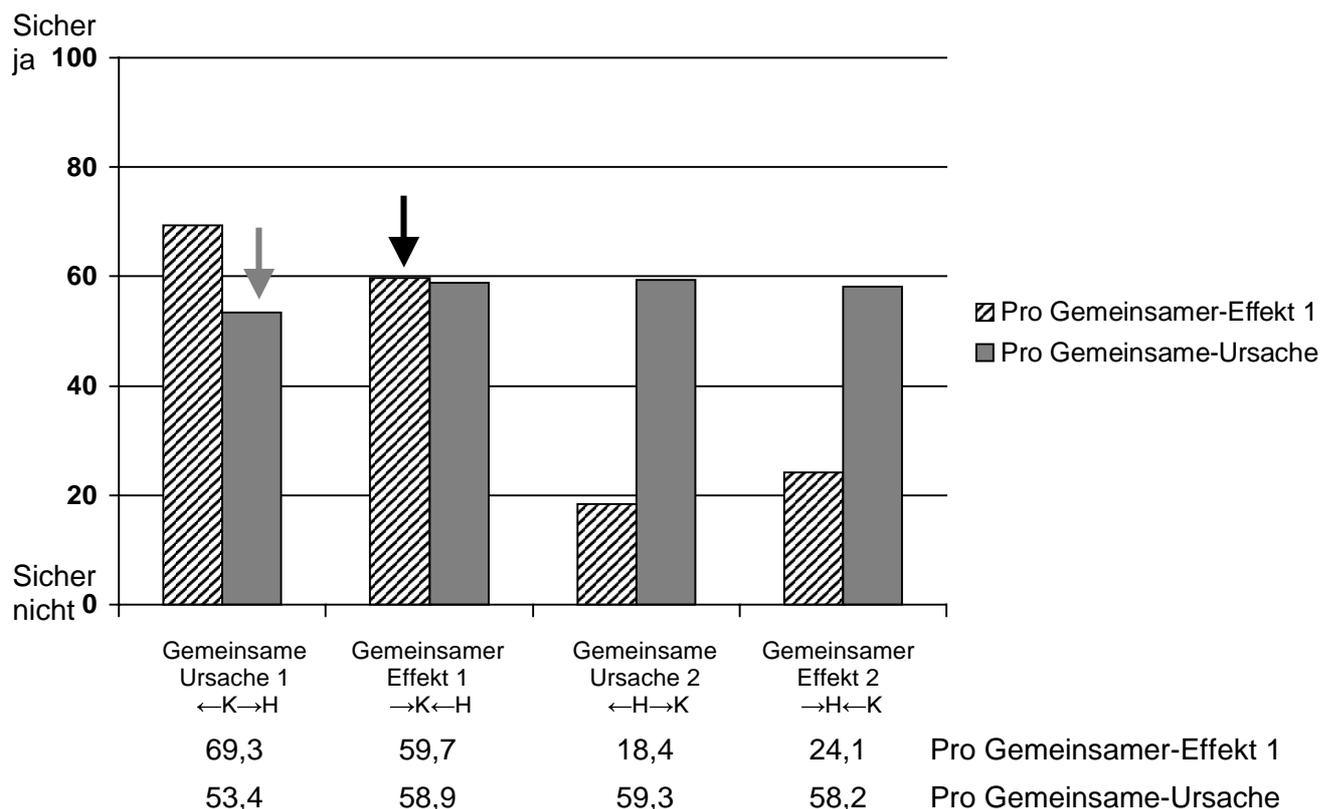


Abb. 3.13: Ergebnisse Experiment 2. Mittelwerte der Beurteilung des Zutreffens der Strukturhypothese. Die beiden Pfeile kennzeichnen das Urteil für die normativ korrekte Strukturhypothese.

Eine Varianzanalyse mit Strukturhypothese als Messwiederholungsvariable und Datensatz als Zwischensubjektfaktor erbrachte signifikante Haupteffekte sowohl für die Strukturhypothese als auch für die Datensätze, $F_{Struktur}(3,138) = 19,2$, $p < 0,01$, $MSE = 348,8$, $F_{Daten}(1,46) = 4,87$, $p < 0,05$, $MSE = 2097,6$. Ebenso wurde die Interaktion signifikant, $F(3,138) = 25,7$, $p < 0,01$, $MSE = 348,8$. Dieses Ergebnis spricht dafür, dass die Probanden systematisch zwischen den Strukturhypothesen und den gezeigten Daten unterschieden und beides jeweils aufeinander bezogen. Wie die Abbildung deutlich macht, sind die beiden Haupteffekte eine Folge der starken Interaktion zwischen den gezeigten Daten und der vorgegebenen Strukturhypothese.

Wie sind die Probanden zu ihrem Urteil gekommen? Eine Analyse des Antwortmusters spricht dafür, dass nicht auf strukturelle Implikationen geachtet wurde. Weder erhielt bei den Daten ‚Pro Gemeinsamer-Effekt 1‘ die gleichnamige Hypothese die höchsten Urteile, noch wurden die beiden Gemeinsame-Ursache Hypothesen bei den Daten, die sie bestätigten, bevorzugt. Das Ergebnis spricht stattdessen dafür, dass nur das Vorliegen der vermuteten Kausalzusammenhänge geprüft wurde. Wurden die Zusammenhänge durch entsprechende Kovariationen in den Daten bestätigt, so ergab sich eine durchweg hohe Zustimmung. Dies trifft bei den Daten ‚Pro Gemeinsame-Ursache‘ auf alle vier Hypothesen zu, hier waren alle drei Ereignisse über gleich starke statistische

Beziehungen miteinander verbunden. Bei den Daten ‚Pro Gemeinsamer-Effekt 1‘ lagen aber nur die vermuteten Zusammenhänge bei den beiden Strukturhypothesen mit der Kennziffer 1 vor. Diese wurden auch angenommen. Bei den beiden Strukturhypothesen mit der Kennziffer 2 wurde ein Zusammenhang zwischen der Ionisierung und der Halogenisierung vermutet. Eine solche Beziehung war in den Daten ‚Pro Gemeinsamer-Effekt 1‘ nicht vorhanden und die Hypothesen wurden zurückgewiesen.

Dafür, dass strukturelle Implikationen keine Rolle für das Urteil spielten, spricht auch eine separate Varianzanalyse, die nur für die Gemeinsame-Ursache- und die Gemeinsamer-Effekt Strukturhypothese 1 durchgeführt wurde. Bei diesen beiden Hypothesen wurden jeweils die beiden vermuteten Kausalzusammenhänge bei beiden Datensätzen bestätigt. Eine Analyse lieferte keine signifikanten Haupteffekte, aber eine signifikante Interaktion, $F(1,46) = 4,54$, $p < 0,05$, $MSE = 306,5$. Eine Interaktion wäre in der Tat zu erwarten gewesen, allerdings, wie die Graphik zeigt, in entgegengesetzter Richtung. Die Gemeinsame-Ursache Strukturhypothese wurde bei den Daten, die gegen sie sprechen (Pro Gemeinsamer-Effekt 1) eher angenommen ($M = 69,3$) als bei ihren eigenen Daten ($M = 53,4$). Bei der Gemeinsamer-Effekt Strukturhypothese trat dieser Unterschied nicht auf. Wahrscheinlich ist dieser Effekt auf Zufall zurückzuführen. Die Richtung des Unterschieds spricht jedenfalls entschieden gegen jedwede Sensitivität für die Verletzung bzw. Bestätigung struktureller Implikationen.

Auch die an das Experiment anschließende Befragung der Probanden erbrachte keinen Hinweis darauf, dass strukturelle Implikationen beachtet worden wären. Vielmehr behaupteten die Probanden, die einzelnen Beobachtungen danach beurteilt zu haben, ob sie für das Vorliegen der beiden Kausalzusammenhänge sprächen oder nicht.

Zusammenfassend ist zu sagen, dass die Ergebnisse eindeutig dagegen sprechen, dass Personen strukturelle Implikationen als Kriterium beim Prüfen von Hypothesen einsetzen.

3.2.2 Experiment 3

Die Befunde des ersten Experimentes dieser Reihe machen deutlich, dass die Probanden nicht auf strukturelle Implikationen geachtet haben. Dies könnte verschiedene Gründe gehabt haben. So könnten die Versuchsteilnehmer die Implikationen verschiedener Kausalstrukturen nicht kennen. Dies hieße, dass ihnen das einzige Kriterium anhand dessen sich verschiedene Kausalstrukturen valide unterscheiden lassen, nicht bekannt wäre. Andererseits könnte die sukzessive Beurteilung mehrerer ähnlicher Strukturhypothesen im ersten Experiment nahegelegt haben, dass mehrere Hypothesen mit den Daten vereinbar sind. Um die Probanden stärker auf die Unterschiede zwischen den Kausalstrukturen aufmerksam zu machen, wurde daher ein zweites Experiment durchgeführt. Aufgabe war es, zwei kausale Strukturhypothesen vergleichend anhand eines Datensatzes zu prüfen. Die Probanden sollten sich also zwischen zwei Hypothesen entscheiden. Hierzu wurden die Gemeinsame-Ursache Hypothese 1 und die Gemeinsamer-Effekt Hypothese 1 aus dem ersten Experiment erneut verwendet. Diese beiden Strukturen postulieren Kausalzusammenhänge zwischen denselben Ereignissen. Daher reicht es für eine Entscheidung zwischen ihnen nicht aus, das Vorhandensein der vermuteten Kausalzusammenhänge zu prüfen. Nur die Berücksichtigung des Zusammenhangs zwischen den beiden nicht kausal verbundenen Ereignissen lässt eine Entscheidung zu.

Als Faktoren wurden einerseits die gezeigten Daten und andererseits der Darbietungs- und Bearbeitungsmodus der Daten variiert. Als Daten kamen wiederum die beiden Datensätze ‚Pro Gemeinsame-Ursache‘ und ‚Pro Gemeinsamer-Effekt‘ aus Experiment eins zum Einsatz. Diese sprechen jeweils eindeutig für das eine und gegen das andere Modell.

Als Darbietungsmodi wurde einmal eine Präsentation im Listenformat gewählt. Damit einher ging die Erlaubnis, sich Notizen zu machen und eine freie Zeiteinteilung. Diese Art der Datenbearbeitung entspricht weitgehend der eines Forschers, der beliebige Zeit und Möglichkeiten hat, sich mit den Daten auseinander zu setzen. Dem entgegengesetzt wurde eine Trial-by-Trial Lernprozedur verwendet, bei der die Daten als graphische Bilder in zufälliger Reihenfolge einmal dargeboten wurden. Notizen waren hierbei nicht erlaubt. Diese Situation entspricht der eines Beobachters in der Welt, welcher nicht die Möglichkeiten hat, sich seine Aufgabe durch externe Hilfen zu vereinfachen. Abbildung 3.14 stellt die vier resultierenden Bedingungen dar.

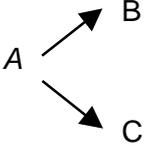
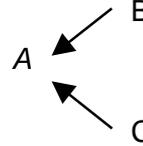
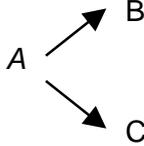
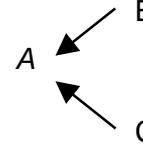
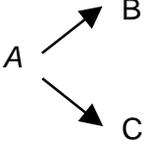
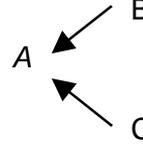
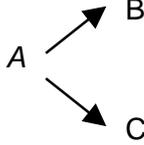
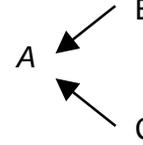
	Darbietung als Liste	Darbietung in Einzelbildern
Daten pro Gemeinsame- Ursache 	 vs. 	 vs. 
Daten pro Gemeinsamer- Effekt 	 vs. 	 vs. 

Abb. 3.14: Versuchsbedingungen Experiment 3

Es wurde vermutet, dass die vergleichende Prüfung der beiden Strukturhypothesen die Unterschiede zwischen den Hypothesen salienter macht. Die Probanden sollten daher bei dieser Aufgabe die unterschiedlichen strukturellen Implikationen der beiden Modelle stärker berücksichtigen. Die Gemeinsame-Ursache Strukturhypothese sollte also bei den Daten ‚Pro Gemeinsame-Ursache‘ und die Gemeinsamer-Effekt Hypothese bei den Daten ‚Pro Gemeinsamer-Effekt‘ bevorzugt werden.

Die Präsentation in Form von Einzelbildern sollte die Beurteilung der Hypothesen erschweren, da die Anforderungen an das Arbeitsgedächtnis erheblich höher sind. So muss eine Vielzahl an Einzelfällen memoriert und daraus Schlussfolgerungen über die Zusammenhänge zwischen den Ereignissen abgeleitet werden. Die Ergebnisse der Datenanalyse müssen dann zu beiden Strukturhypothesen gleichzeitig in Bezug gesetzt werden. Es erscheint unwahrscheinlich, dass Probanden hierzu in der Lage sind. Bei der Präsentation in Listenform sollte dies kein Problem sein, da die Daten stets alle zur Verfügung stehen und beliebig bearbeitet und verrechnet werden können.

Neben der Beurteilung der Hypothese wurde zusätzlich erfasst, ob die Probanden den Zusammenhang zwischen den nicht direkt kausal verbundenen Ereignissen richtig erfassen. Das Wissen über diese Relation ist entscheidend, damit strukturelle Implikationen berücksichtigt werden können.

Probanden und Design

40 Studierende der Psychologie an der Universität Göttingen nahmen teil. Alle Probanden befanden sich in den ersten beiden Semestern ihres Studiums und hatten noch nicht an einer Veranstaltung zur Versuchsplanung teilgenommen. Sie erhielten für ihre Teilnahme eine Versuchspersonenstunde bescheinigt. Die Probanden wurden per Zufall

einer der vier Versuchsbedingungen zugeteilt. Das Experiment wurde im Einzelversuch durchgeführt.

Versuchsdurchführung und Material

Alle Probanden erhielten dieselbe Aufgabe. Sie wurden instruiert, dass der Thalamus die zentrale cortikale Schaltstelle im Gehirn sei. Seine Kerne würden über unterschiedliche Bahnen aktiviert und würden ihre Aktivität an verschiedene Regionen des Cortex weitergeben. Wenig wisse man bisher, wie die Kerne des Thalamus sich untereinander beeinflussen. Bekannt sei nur, dass es kausale Zusammenhänge geben müsse. Nun sei es Neurophysiologen erstmals gelungen, mit Hilfe der Positronen-Emissions-Tomographie (PET) Bilder von der Aktivität der einzelnen Kerne zu machen. Diese Bilderserie solle dazu dienen, die Kausalzusammenhänge von drei der wichtigsten Kerne aufzuklären. Die drei Kerne seien der Pulvinar, welcher mit dem retikulären System in Verbindung stehe, der vordere ventroorale Kern, welcher mit dem motorischen System zusammenhänge und der Nucleus anterior, der mit dem limbischen System zusammenarbeite. Des Weiteren wurden die Probanden informiert, dass die beteiligten Wissenschaftler zwei konkurrierende Hypothesen darüber verträten, wie sich die Kerne kausal beeinflussen. *Hypothese 1* sei, dass die Aktivität im Pulvinar sowohl die Aktivität im Nucleus ventrooralis als auch die Aktivität im Nucleus anterior beeinflusse. *Die Konkurrenzhypothese 2* sei, dass sowohl die Aktivität im Nucleus ventrooralis als auch die Aktivität im Nucleus anterior die Aktivität im Pulvinar beeinflusse.⁴⁴ Die Aufgabe der Probanden sei es nun, mit Hilfe der ermittelten Daten zwischen den beiden Hypothesen zu entscheiden. Sie würden nach ihrer Analyse der Daten gefragt werden, mit welcher Wahrscheinlichkeit die beiden Hypothesen zuträfen.

Die Probanden wurden gebeten, die Instruktion und ihre Aufgabe mit eigenen Worten wiederzugeben, um das Instruktionsverständnis zu überprüfen. Etwaige Unklarheiten wurden beseitigt. Danach erhielten die Teilnehmer in der Listenbedingung eine zweiseitige Liste mit 96 Beobachtungsfällen. In der ersten Spalte waren die Beobachtungen durchnummeriert, in den folgenden Spalten war eingetragen, welche der drei Thalamuskern aktiv oder inaktiv waren. Zusätzlich erhielten die Probanden in der Listenbedingung Stift und Notizpapier. In den beiden anderen Bedingungen bekamen die Probanden sechsundneunzig fingierte PET-Scan Aufnahmen der Aktivität in den Thalamuskern ausgehändigt. Die untenstehende Abbildung 3.15 zeigt ein Beispiel. Die Originalbilder waren farbig, aktive Areale waren rot, inaktive blau gefärbt.

⁴⁴ In den gängigen Lehrbüchern zur Biopsychologie (z.B. Birbaumer & Schmidt, 1990) werden nur die Kerne des Thalamus, aber nicht ihre Verbindung untereinander besprochen. Es ist daher davon auszugehen, dass die Probanden keinerlei Vorwissen über die angeführten Kausalzusammenhänge hatten.

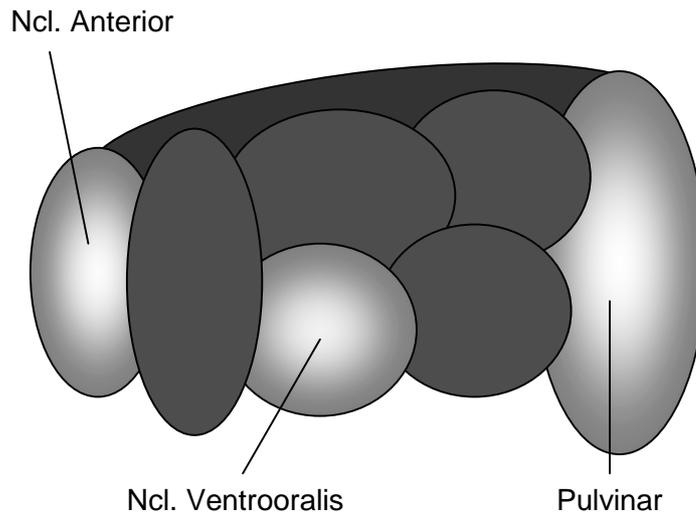


Abb. 3.15: Versuchsmaterial Experiment 3: PET-Scan Aufnahmen des Thalamus.

Die Probanden durften die Aufnahmen einmal durchgehen, Zurückblättern oder Ordnen war nicht erlaubt. Sie wurden zudem darauf hingewiesen, dass die Reihenfolge der Bilder nicht den Verlauf der Aktivität wiedergebe, die Abfolge sei zufällig.

Nach der Analyse der Daten wurde den Probanden ein Ratingbogen vorgelegt, auf dem beide Strukturhypothesen erneut abgebildet und durch eine offene Skala miteinander verbunden waren. Sie wurden gebeten einzuschätzen, welche der beiden Hypothesen eher zutrifft. Wenn sie sicher seien, dass eine der Hypothesen zuträfe, sollten sie die Skala bei der entsprechenden Hypothese ganz außen ankreuzen. Die untenstehende Abbildung zeigt verkleinert die Ratingskala. Die offene Skala war im Original 10 cm lang.

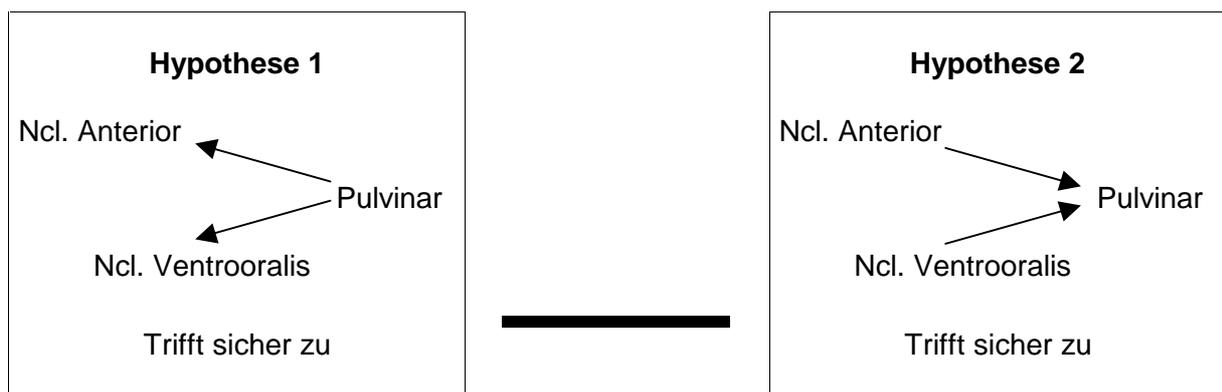


Abb. 3.16: Ratingskala Experiment 3
(verkleinerte Abbildung, Original in A4 Querformat)

Die Probanden erhielten keine Rückmeldung zu ihrem Urteil. Sie wurden dann gebeten, ihr Vorgehen zu erläutern und ihre Einschätzung zu begründen. Daraufhin sollten sie die Korrelation zwischen der Aktivität im Nukleus anterior und der Aktivität im Nukleus ventrooralis auf einer Skala zwischen -1 („Wenn der eine aktiv ist, ist der andere nicht aktiv“)

bis +1 („wenn der eine aktiv ist, dann ist auch der andere aktiv“) einschätzen. Das Maß der Korrelation war allen Teilnehmenden bekannt.

Präsentierte Daten

Die Daten wurden von Experiment 2 übernommen. Sie sprachen entweder eindeutig für die Gemeinsame-Ursache Hypothese oder für die Gemeinsamer-Effekt Hypothese. Es wurde lediglich die Bezeichnung der einzelnen Ereignisse ausgetauscht.

Ergebnisse

Die Einschätzungen auf der offenen Skala wurden durch Ausmessen auf eine Skala von -5 (Gemeinsame-Ursache Strukturhypothese trifft sicher zu) bis +5 (Gemeinsamer-Effekt Strukturhypothese trifft sicher zu) übertragen. Abbildung 3.17 gibt in der linken Hälfte die Mittelwerte für die Hypothesenbeurteilung und in der rechten Hälfte die Mittelwerte der Korrelationsschätzung an. Ein Proband weigerte sich in der Pro Gemeinsame-Ursache Bedingung bei Darbietung von Einzelbildern eine Einschätzung der Korrelation abzugeben. Dieser Mittelwert beruht daher nur auf neun Probanden, alle anderen auf 10 Probanden.

Hypothesenbeurteilung (Skala: -5 „Gemeinsamer-Effekt Hypothese trifft zu“ - +5 „Gemeinsame-Ursache Hypothese trifft zu“)			Geschätzte Korrelation Aktivität im Ncl. anterior und Ncl. ventrooralis (Skala: -1 - +1)		
	Darbietung als Liste	Darbietung in Einzelbildern		Darbietung als Liste	Darbietung in Einzelbildern
Pro Gemeinsame Ursache	-0,7	0,3	Pro Gemeinsame Ursache	0,33	0,33
Pro Gemeinsamer Effekt	1,2	1,1	Pro Gemeinsamer Effekt	0,10	0,07

Abbildung 3.17: Ergebnisse Experiment 3

Für die Beurteilung der Hypothesen ergaben sich Mittelwerte, die in allen Bedingungen um Null herum lagen. Eine Varianzanalyse mit Darbietungsmodus und Daten als Zwischensubjektfaktoren erbrachte keine signifikanten Effekte, $F_{\text{Daten}}(1,36) = 1,62$, $p = 0,21$, $F_{\text{Darbietung}}(1,36) = 0,18$, $p = 0,67$, $F_{\text{Interaktion}}(1,36) = 0,27$, $p = 0,61$, $MSE = 11,24$. Eine Betrachtung der Einzelantworten zeigte, dass sich in allen vier Bedingungen eine trimodale Verteilung ergab. Je ein Drittel der Probanden entschied sich unabhängig von den Daten und der Präsentationsform der Daten für die Gemeinsame-Ursache Hypothese, die Gemeinsamer-Effekt Hypothese oder gab keiner der beiden Hypothesen den Vorzug.

Bei der Einschätzung der Korrelation zeigte sich dem gegenüber ein klarer Effekt. Bei den Daten ‚Pro Gemeinsame Ursache‘ wurde der Zusammenhang der Aktivität in den beiden nicht kausal verbundenen Ereignissen deutlich höher eingeschätzt als bei den Daten ‚Pro Gemeinsamer-Effekt‘. Eine Varianzanalyse erbrachte einen signifikanten Haupteffekt für die verschiedenen Datensätze $F_{Daten}(1,35) = 9,86, p < 0,01, MSE = 0,06$. Kein Effekt zeigte sich für die Präsentation der Daten ($F < 1$). Ebenso war die Interaktion weit von jeder Signifikanz entfernt ($F < 1$).

Die Ergebnisse widerlegen klar die angeführten Hypothesen. Weder führte der direkte Vergleich der beiden Hypothesen zu einer erhöhten Berücksichtigung der strukturellen Implikationen, noch führte die Manipulation der Datendarbietung zu einem veränderten Antwortverhalten.⁴⁵

Die Beurteilungen der Hypothese replizieren den Befund von Experiment 2. Es zeigte sich keine systematische, von den Daten abhängige Präferenz für eine der beiden Hypothesen. Insbesondere, dass sich unabhängig von den Daten genauso viele Probanden für die Gemeinsame-Ursache wie für die Gemeinsamer-Effekt Strukturhypothese entschieden, spricht klar gegen eine Nutzung von strukturellen Implikationen als Prüfkriterium. Dabei verfügten die Probanden aber über das für die richtige Entscheidung relevante Wissen. Wie die Einschätzungen der Korrelationen zeigen, erkannten die Probanden den Zusammenhang der beiden nicht kausal verbundenen Ereignisse. Trotzdem nutzten sie dieses Wissen nicht. In den verbalen Begründungen für ihr Urteil bezogen sich die Probanden häufig darauf, dass die Daten die vermuteten Kausalzusammenhänge des jeweils bevorzugten Modells bestätigten. Dabei waren sich die Probanden scheinbar nicht bewusst, dass dies ebenso für das konkurrierende Modell der Fall war.

⁴⁵ Die Ergebnisse sind aufgrund der geringen Versuchspersonenzahl mit Vorsicht zu betrachten. Ihre Aussagekraft erhalten sie erst in Bezug auf die Ergebnisse der anderen Experimente, die zu dieser Fragestellung durchgeführt wurden. Hierbei zeigen sich konvergierende Ergebnisse. Dies spricht für die Zuverlässigkeit des Befundes.

3.2.3 Experiment 4

Die Befunde der ersten beiden Experimente dieser Reihe sprechen dafür, dass strukturelle Implikationen bei der Prüfung von Strukturhypothesen vernachlässigt werden. Ein mögliches Argument gegen diese Schlussfolgerung mag sein, dass die Daten in diesen Experimenten nicht normativ waren. Wenn die Probanden extrem sensitiv gegenüber den von einer Struktur implizierten bedingten wie unbedingten Abhängigkeiten wären, dann könnten unter Umständen solche nicht-normativen Daten die Versuchsteilnehmer so verwirren, dass sie auf eine andere Teststrategie umstiegen und dann nur das Vorhandensein der vermuteten Kausalzusammenhänge überprüften. Gemäß dieser These versuchten die Probanden in den beiden Experimenten also strukturelle Implikationen zu testen. Da aber die Verteilung der Beobachtungen nicht normativ war, setzten sie deshalb eine nicht normative Heuristik zur Lösung der Aufgabe ein. Um diesen Einwand auszuschließen, war das Ziel von Experiment 4, die Befunde der beiden vorhergehenden Experimente auch für normativ richtige Datensätze zu replizieren. Die Aufgabe der Probanden war es, nur eine kausale Strukturhypothese anhand eines Datensatzes zu überprüfen. Als Hypothesen wurden die aus Experiment 3 bekannte Gemeinsame-Ursache und Gemeinsamer-Effekt Strukturhypothese erneut verwendet. Beide Hypothesen gingen wie zuvor von Kausalzusammenhänge zwischen denselben Ereignissen aus. Zu jeder dieser beiden Hypothesen wurde ein entsprechender, sie vollständig bestätigender Datensatz entwickelt. Um normativ richtige Daten zu gewährleisten, wurden diese mittels eines Bayesianischen Netzwerks, welches eine Gemeinsame-Ursache oder eine Gemeinsame-Effekt Struktur aufwies, generiert. Dies führte dazu, dass die Daten sämtliche unbedingten und bedingten implizierten Abhängigkeiten des jeweiligen Modells vollständig aufwiesen und auch die Verteilung der verschiedenen Beobachtungsfälle einer Bayesianischen Analyse standhielt.

Was ist für dieses Experiment zu erwarten? Wenn Versuchsteilnehmer tatsächlich extrem sensitiv für strukturelle Implikationen sind, dann sollte sich diese Sensitivität hier zeigen. Dennoch erscheint dies unwahrscheinlich. In keinem der Vorgängerexperimente gab es einen Hinweis darauf, dass die Implikationen der Strukturmodelle genutzt wurden. Weder wurden implizierte Abhängigkeiten zwischen den nicht direkt kausal verbundenen Ereignissen überprüft noch implizierte Ereignisverteilungen. Es ist daher zu erwarten, dass die Veränderung der Datensätze die Prüfurteile der Probanden nicht verbessert. Sie sollten vielmehr unabhängig von den gezeigten Daten beide Hypothesen annehmen, da die vermuteten Kausalzusammenhänge bestätigt werden. Abbildung 3.18 fasst die Bedingungen und Vorhersagen zusammen.

	Gemeinsame-Ursache Hypothese		Gemeinsamer-Effekt Hypothese	
	Prüfung struktureller Implikationen	Prüfung vermuteter Kausalzshg.	Prüfung struktureller Implikationen	Prüfung vermuteter Kausalzshg.
Daten pro Gemeinsamer-Effekt	↓	↑	↑	↑
Daten pro Gemeinsame-Ursache	↑	↑	↓	↑

Abb. 3.18: Versuchsbedingungen und Vorhersagen Experiment 4, ↑ bedeutet Zustimmung, ↓ bedeutet Ablehnung

Probanden und Design

48 Studierende verschiedener Fachbereiche der Universität Göttingen nahmen an dem Experiment teil. Sie wurden über Aushänge angeworben und erhielten DM 10,-- als Bezahlung. Keiner der Probanden hatte zuvor an einem anderen Experiment der Serie teilgenommen. Die Versuchsteilnehmer wurden einer der vier Versuchsbedingungen per Zufall zugeteilt. Das Experiment wurde im Einzelversuch durchgeführt.

Versuchsdurchführung und Material

Alle Probanden erhielten die Aufgabe, eine kausale Strukturhypothese anhand eines Datensatzes zu überprüfen. Wie in Experiment 2 bezogen sich die Hypothesen auf den Zusammenhang zwischen drei Kernen des Thalamus.

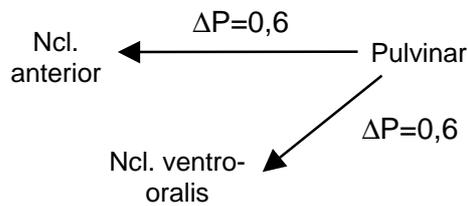
Die Probanden wurden instruiert, dass der Thalamus die zentrale cortikale Schaltstelle im Gehirn sei und hier Informationen aus allen Regionen des Nervensystems zusammenkämen. Der Thalamus selbst bestehe aus einer Vielzahl unabhängiger Kerne. Man wisse bisher wenig darüber, wie die Kerne des Thalamus sich untereinander beeinflussen. Bekannt sei nur, dass es kausale Zusammenhänge geben müsse. Nun sei es Neurophysiologen erstmals gelungen, mit Hilfe der Positronen-Emissions-Tomographie (PET) Bilder von der Aktivität der einzelnen Kerne zu machen. Diese Bilderserie solle dazu dienen, die Kausalzusammenhänge von drei der wichtigsten Kerne aufzuklären. Die drei Kerne seien der Pulvinar, der vordere ventroorale Kern und der Nucleus anterior. Die Wissenschaftler hätten eine Hypothese darüber, wie sich die Kerne kausal beeinflussen. In

der Gemeinsame-Ursache Bedingung lasen die Probanden daraufhin, die Hypothese sei, dass die Aktivität im Pulvinar die Aktivität im Nucleus ventrooralis und die Aktivität im Nucleus anterior beeinflusse. Der Pulvinar solle sich also auf die beiden anderen Areale auswirken. In der Gemeinsamer-Effekt Bedingung lautete die Hypothese, dass die Aktivität im Nucleus ventrooralis und die Aktivität im Nucleus anterior die Aktivität im Pulvinar beeinflusse. Die beiden Nuklei sollten sich also auf den Pulvinar auswirken. Ihre Aufgabe sei es, mit Hilfe der ermittelten Daten die Hypothese zu überprüfen. Sie würden später gefragt werden, ob die Hypothese zutreffe.

Im Anschluss an die Instruktion wurden die Probanden gebeten, die Strukturhypothese aufzuzeichnen. Diese Aufgabe diente dazu, das Instruktionsverständnis zu überprüfen. Etwaige Unklarheiten auf Seiten der Versuchsteilnehmer wurden an dieser Stelle beseitigt. Danach erhielten die Probanden entweder die Daten ‚Pro Gemeinsame-Ursache‘ oder ‚Pro Gemeinsamer-Effekt‘. Die Daten wurden in Form von 100 Karten mit fingierten PET Scan Aufnahmen vorgegeben. Die Abbildungen wurden dabei von Experiment 2 übernommen. Die Bedeutung der Farben (rot = aktiv, blau = inaktiv) wurde den Probanden anhand eines Beispiels erläutert. Ebenso wurden sie darauf hingewiesen, dass die Reihenfolge der Karten zufällig sei. Die weitere Bearbeitung der Daten war den Probanden freigestellt, das Ordnen der Karten und Notizenmachen waren erlaubt. Daran anschließend wurden sie gefragt, wie wahrscheinlich es sei, dass die Hypothese zutreffe. Sie sollten ihre Einschätzung auf einer Skala von 0 (Hypothese trifft nicht [zu 0%] zu) bis 100 (Hypothese trifft sicher [zu 100%] zu) angeben.

Präsentierte Daten

Die Daten wurden über zwei Bayesianische Netzwerke generiert. Abbildung 3.19 zeigt diese beiden Netzwerke mit den zugehörigen Formeln und Wahrscheinlichkeiten. Wie bei den Daten der beiden ersten Experimente wurden die bedingten Wahrscheinlichkeiten der Effekte bei An- und Abwesenheit der Ursache in beiden Modellen gleichgesetzt. Daher waren die für den tatsächlichen Kausalzusammenhang indikativen Kontingenzen bei beiden Modellen identisch. Das heißt bei der Gemeinsame-Ursache Struktur war die unbedingte Kontingenz der Ursache auf ihre beiden Effekte $\Delta P = 0,6$ und bei der Gemeinsamer-Effekt Struktur war die auf die Abwesenheit der zweiten Ursache bedingte Kontingenz $\Delta P_C = 0,6$. Zusätzlich wurde die Basisrate der Ursachen in beiden Modellen auf $P = 0,5$ gesetzt.

Gemeinsame-Ursache Struktur

Bayesianisches Modell

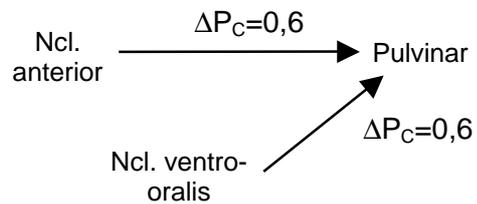
$$P(\text{PUL.}, \text{VEN.}, \text{ANT.}) = P(\text{ANT.}|\text{PUL.}) \cdot P(\text{VEN.}|\text{PUL.}) \cdot P(\text{PUL.})$$

Gesetzt wurde

$$P(\text{pul.}) = P(\sim\text{pul.}) = 0,5$$

$$P(\text{ant.}|\text{pul.}) = P(\text{ven.}|\text{pul.}) = 0,8$$

$$P(\text{ant.}|\sim\text{pul.}) = P(\text{ven.}|\sim\text{pul.}) = 0,2$$

Gemeinsamer-Effekt Struktur

Bayesianisches Modell

$$P(\text{PUL.}, \text{VEN.}, \text{ANT.}) = [1 - (1 - P(\text{PUL.}|\text{ANT.})) \cdot (1 - P(\text{PUL.}|\text{VEN.}))] \cdot P(\text{ANT.}) \cdot P(\text{VEN.})$$

Gesetzt wurde

$$P(\text{ant.}) = P(\sim\text{ant.}) = P(\text{ven.}) = P(\sim\text{ven.}) = 0,5$$

$$P(\text{pul}|\text{ant.}\sim\text{ven.}) = P(\text{pul}|\text{ven.}\sim\text{ant.}) = 0,8$$

$$P(\text{pul}|\sim\text{ven.}\sim\text{ant.}) = 0,2$$

Abb. 3.19: Konstruktionsprinzipien der Daten für Experiment 4. Die Bayesianischen Modelle beziehen sich auf Ereignisvariablen (Großbuchstaben), die gesetzten Wahrscheinlichkeiten auf Ereignisausprägungen (Kleinbuchstaben)

Die Verteilung der Ereigniskombinationen ergibt sich dann über das Bayesianische Modell multipliziert mit den 100 Beobachtungsfällen. Abbildung 3.20 zeigt die resultierenden Häufigkeiten. Diese Daten erfüllen alle strukturellen Implikationen der beiden Modelle.

Daten Pro Gemeinsame-Ursache

	ant.		~ant.	
	ven.	~ven.	ven.	~ven.
pul	32	8	8	2
~pul.	2	8	8	32

Daten Pro Gemeinsamer-Effekt

	ant.		~ant.	
	ven.	~ven.	ven.	~ven.
pul	24	20	20	5
~pul.	1	5	5	20

Abb.3.20: Daten zu Experiment 4

Welche Zusammenhänge finden die Probanden, wenn sie diese Daten analysieren? Abbildung 3.21 fasst die Antworten zusammen. Die Werte entlang der gestrichelten Linien geben die Kontingenzen an, die sich ergeben, wenn der Zusammenhang aufgrund der entgegengesetzten Hypothese in der anderen Richtung analysiert wird. Prüfen wir beispielsweise eine Gemeinsamer-Effekt Hypothese anhand der Daten ‚Pro Gemeinsame-Ursache‘, so sollten wir die Kontingenz des Pulvinar auf den Nucleus anterior bedingt auf die Inaktivität des Nucleus ventrooralis bestimmen. Tun wir dies, so finden wir eine bedingte

Kontingenz von $\Delta P_C = 0,44$. Nehmen wir fälschlicherweise die unbedingte Kontingenz, so erhalten wir $\Delta P = 0,60$. Wie aus der Abbildung ersichtlich, besteht eine Kontingenz von $\Delta P = 0,36$ zwischen den beiden Effekten bei den Daten ‚Pro Gemeinsame-Ursache‘. Dies ist auch die Zusammenhangsstärke, die von der Struktur impliziert wird. Bei den Daten ‚Pro Gemeinsamer-Effekt‘ sind die Ursachen korrekterweise unabhängig.

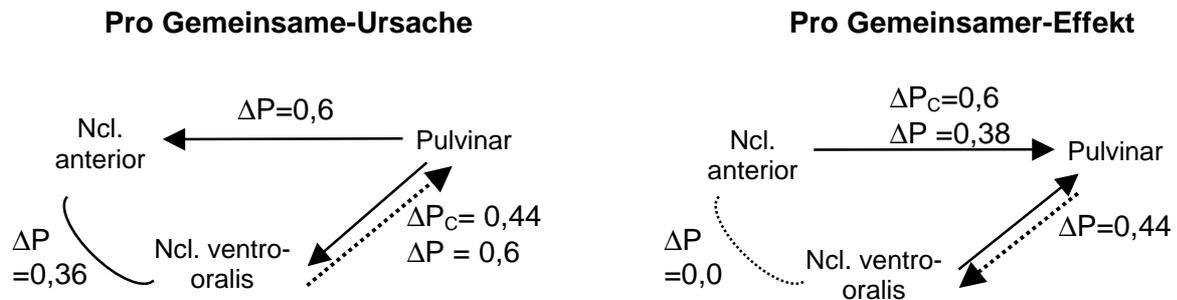


Abb. 3.21: Zusammenhänge zwischen den Ereignissen in Experiment 4

Ergebnisse

Die Mittelwerte der Prüfurteile in den vier Bedingungen sind in Abbildung 3.22 dargestellt.

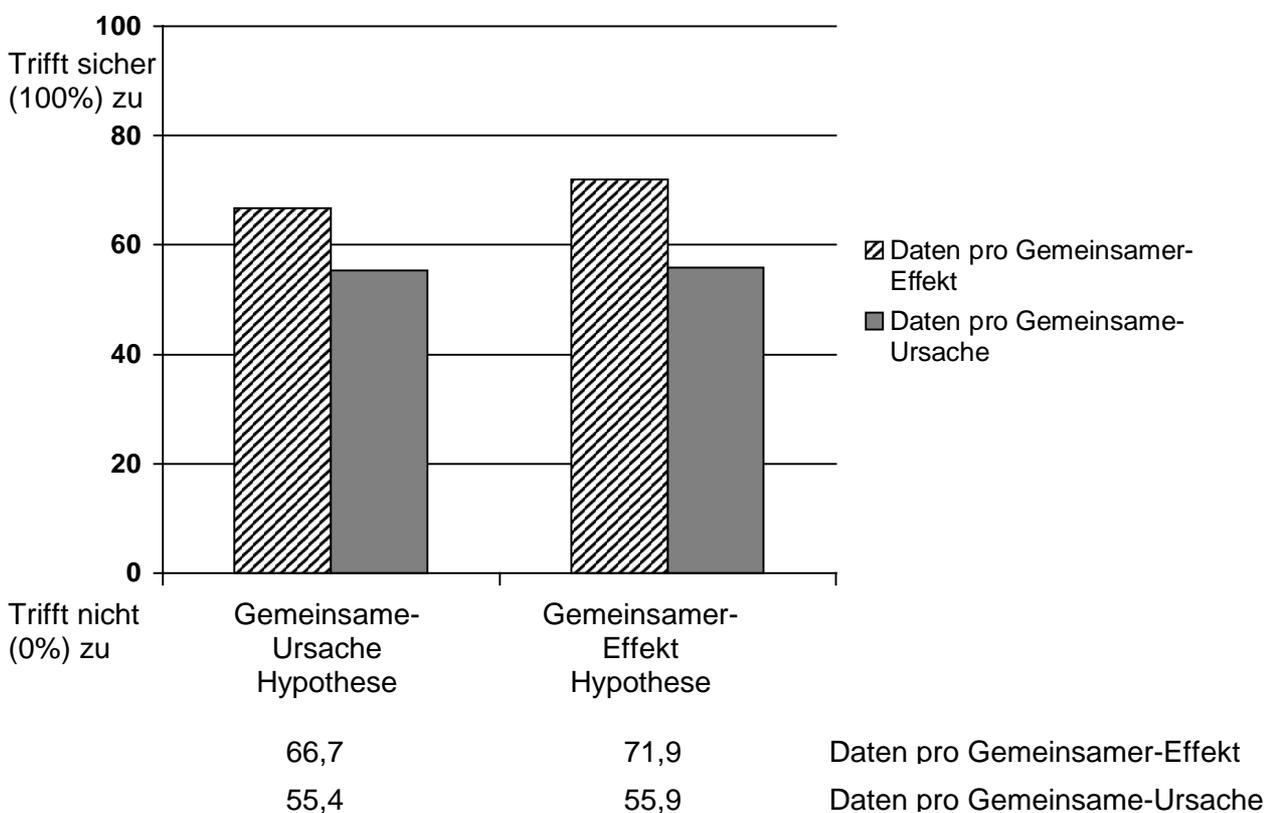


Abb. 3.22: Ergebnisse Experiment 4: Mittelwerte der Prüfurteile

Eine Varianzanalyse mit Strukturhypothese und Datensatz als Zwischensubjektfaktoren erbrachte einen signifikanten Haupteffekt für den gezeigten Datensatz, $F(1,44) = 4,63$, $p < 0,05$, $MSE = 481,0$. Ansonsten zeigte sich weder ein Haupteffekt der Strukturhypothese noch eine Interaktion (jeweils $F < 1,0$).

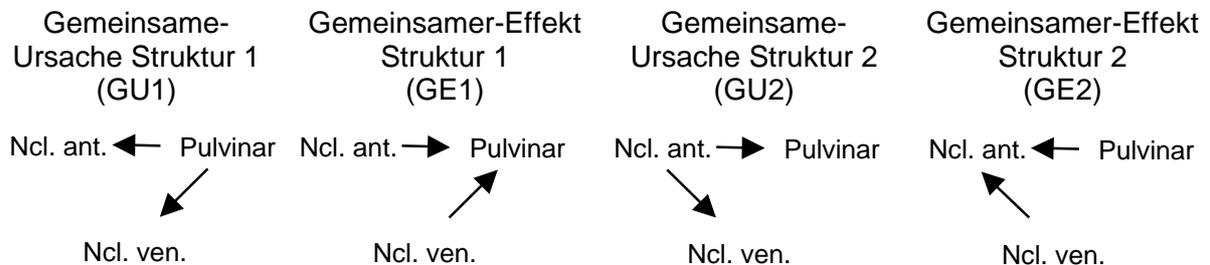
Dass es einen signifikanten Unterschied zwischen den Datensätzen gibt, ist überraschend. Wie oben ausgeführt sind beide Datensätze weitestgehend parallelisiert. Eine Analyse der Notizen der Probanden erbrachte keine Hinweise auf Unterschiede im Vorgehen. Betrachtet man die Verteilung der Urteile, so fällt auf, dass in allen Bedingungen ein bis zwei Probanden die vorgelegte Hypothese ablehnten. Die Urteile dieser Probanden wichen bei den Daten ‚Pro Gemeinsame-Ursache‘ stärker von den positiven Urteilen der anderen Versuchsteilnehmer ab. Die Varianz war in dieser Datenbedingung deshalb auch tendenziell größer ($s_{\text{Gemeinsame-Ursache}} = 23,8$ vs. $s_{\text{Gemeinsamer-Effekt}} = 18,9$). Diese Beobachtung und die Befunde der Vorgängerexperimente legen nahe, dass es sich wahrscheinlich um einen Zufallseffekt handelt.

Das Fehlen einer Interaktion spricht dafür, dass die Probanden nicht überprüfen, ob die strukturellen Implikationen erfüllt sind. Sie machen keine systematischen Unterschiede zwischen den Hypothesen, obwohl nur bei dem jeweils gleichnamigen Datensatz die Daten mit dem Modell vereinbar sind. Damit wird der Effekt der beiden ersten Experimente dieser Serie repliziert. Dieser ist also nicht darauf zurückzuführen, dass die Daten nicht über Bayesianische Netze generiert wurden. Trotz der normativ korrekten Daten in diesem Experiment berücksichtigten die Probanden strukturelle Implikationen nicht.

3.2.4 Experiment 5

Der übereinstimmende Befund der vorherigen drei Experimente war, dass Probanden bei der Prüfung von Strukturhypothesen nicht zwischen Modellen differenzieren, die kausale Zusammenhänge zwischen denselben Ereignissen vermuten. Eine Frage, die sich hieran anschließt, ist, ob es andere Kriterien gibt, mittels derer sie zwischen solchen Strukturhypothesen unterscheiden können. Das Ziel von Experiment 5 ist es herauszufinden, ob die Notwendigkeitsrelation ein solches Kriterium ist. In der philosophischen Literatur (Hume, 1987; Mackie, 1974) werden Ursachen häufig als notwendige Bedingungen konzeptualisiert. Ein Effekt tritt nie auf, wenn nicht zuvor eine Ursache aufgetreten ist. Diese logische Bestimmung der Ursache trifft sich mit unserer Intuition, dass es für jedes Ereignis eine Ursache gibt. Es ist daher plausibel anzunehmen, dass Probanden dieser Notwendigkeitsbeziehung eine besondere Bedeutung beimessen. Sie sollten Ereignisse, welche notwendig für das Eintreten anderer Ereignisse sind, als Ursachen ansehen.

In Experiment 5 hatten die Probanden wie zuvor die Aufgabe, eine Strukturhypothese anhand eines Datensatzes zu überprüfen. Hierfür wurden nochmals alle Strukturen aus Experiment 2 eingesetzt. Als Hintergrundgeschichte diente erneut der Zusammenhang der drei Thalamuskern. Abbildung 3.23 stellt die vier Strukturhypothesen im Überblick dar.

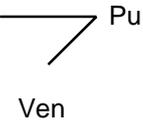
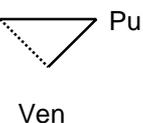
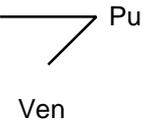
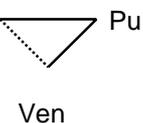


*Abb. 3.23: Strukturhypothesen Experiment 5.
Ncl. ant. = Nukleus anterior, Ncl. ven. = Nukleus ventrooralis*

Als zweiter Faktor wurden wieder die vorgelegten Daten manipuliert. Die beiden Datensätze wurden über ein Bayesianisches Netzwerk mit der Gemeinsame-Ursache Struktur 1 und eines mit der Gemeinsamer-Effekt Struktur 1 erzeugt. Zusätzlich war in den Daten ‚Pro Gemeinsame-Ursache‘ die Aktivität des Pulvinar notwendig für eine Aktivität der beiden anderen Nuklei, während bei den Daten ‚Pro Gemeinsamer-Effekt‘ die Aktivität eines der beiden Nuklei notwendig für eine Aktivität des Pulvinar war. Der Pulvinar war in diesem Fall also niemals aktiv, wenn keiner der beiden anderen Kerne aktiv war. Die beiden Datensätze unterschieden sich ebenfalls in dem Zusammenhang der beiden nicht direkt kausal verbundenen Ereignisse. In den Daten ‚Pro Gemeinsame-Ursache‘ korrelierten die beiden Effekte (die Aktivitäten in den beiden Nuklei) wie von der Struktur impliziert. In den Daten ‚Pro Gemeinsamer-Effekt‘ waren die beiden Ursachen (ebenfalls die Aktivitäten in den beiden Nuklei) unabhängig voneinander.

Experiment 5 testet mit diesen Daten einen seltenen Grenzfall. Es ist eine der einfachsten möglichen Bedingungen für die Prüfung einer Strukturhypothese. Nur drei Ereignisse sind über zwei Kausalrelationen verbunden. Die statistischen Beziehungen sind stark, die strukturellen Implikationen sind entweder voll erfüllt oder nicht erfüllt, und ein Teil der Ereignisse ist notwendig für das Auftreten der anderen. Sofern Probanden überhaupt in der Lage sind, Strukturhypothesen richtig zu prüfen, dann sollte es sich in diesem Experiment zeigen.

Folgende Vorhersagen ergeben sich einerseits für ein normatives Vorgehen unter zu Hilfenahme der Notwendigkeitsbeziehung und andererseits für ein Prüfen der vermuteten Kausalzusammenhänge.

Prüfen struktureller Implikationen und/oder logischer Relationen				
	Gemeinsame- Ursache Struktur 1 Ant ← Pul → Ven	Gemeinsamer- Effekt Struktur 1 Ant → Pul ← Ven	Gemeinsame- Ursache Struktur 2 Ven ← Ant → Pul	Gemeinsamer Effekt Struktur 2 Ven → Ant ← Pul
Pro Gemeinsamer- Effekt Ant  Pul Ven	↓	↑	↓	↓
Pro Gemeinsame- Ursache Ant  Pul Ven	↑	↓	↓	↓
Prüfen der vermuteten Zusammenhänge				
Pro Gemeinsamer- Effekt Ant  Pul Ven	↑	↑	↓	↓
Pro Gemeinsame- Ursache Ant  Pul Ven	↑	↑	→	→

*Abb. 3.24: Vorhersagen Experiment 5;
 ↑ bedeutet Zustimmung, ↓ bedeutet Ablehnung, → steht für ein mittleres Urteil*

In der oberen Hälfte sind die Vorhersagen aufgetragen für den Fall, dass die Probanden normativ vorgehen. In diesem Fall sollten die beiden Strukturhypothesen mit der Kennziffer zwei in beiden Datenbedingungen abgelehnt werden. Dies liegt bei den Daten ‚Pro Gemeinsamer-Effekt‘ darin begründet, dass einer der beiden vermuteten Kausalzusammenhänge fehlt. Bei den Daten ‚Pro Gemeinsame-Ursache‘ ist einer der beiden vermuteten Zusammenhänge (Ncl. anterior - Ncl. ventrooralis) relativ schwach. Dies führt dazu, dass unter der Gemeinsamen-Ursache Hypothese 2 der starke Zusammenhang zwischen der Aktivität im ventrooralen Kern und der Aktivität im Pulvinar nicht impliziert wird. Unter der Gemeinsamer-Effekt Hypothese 2 würden die vermeintlichen Ursachen stärker korrelieren als eine der Ursache mit ihrem Effekt. Dies widerspricht der Hypothese. Für die Hypothesen mit der Kennziffer 1 ist die Vorhersage klar. Bei den Daten, für die das jeweilige

Strukturmodell Pate stand, sollte diese Hypothese angenommen werden. Das jeweils andere Strukturmodell sollte abgelehnt werden.

Eine Analyse der logischen Beziehungen zwischen den Aktivitäten der Kerne führt zu derselben Vorhersage. Bei den Daten ‚Pro Gemeinsame-Ursache‘ ist die Aktivität des Pulvinar notwendig für die Aktivität der beiden anderen Kerne. Dies entspricht der Logik der Gemeinsame-Ursache Hypothese 1. Bei den Daten ‚Pro Gemeinsamer-Effekt‘ ist eine Aktivität des Nukleus ventrooralis oder des Nukleus anterior notwendig für eine Aktivität des Pulvinar. Dies korrespondiert zur Gemeinsamer-Effekt Hypothese 1.

Welche Vorhersagen ergeben sich aber, wenn die Probanden wie in den anderen Experimenten lediglich die Zusammenhänge überprüfen? In diesem Fall sollten bei den Daten ‚Pro Gemeinsamer-Effekt‘ die beiden Hypothesen mit der Kennziffer 1 angenommen werden, da bei diesen die beiden vermuteten Zusammenhänge vorliegen. Die anderen beiden Hypothesen sollten zurückgewiesen werden. Bei den Daten ‚Pro Gemeinsame-Ursache‘ sollte ebenfalls den Hypothesen mit der Kennziffer 1 zugestimmt werden. Die beiden Hypothesen mit der Kennziffer 2 sollten etwas niedriger beurteilt werden, da einer der beiden Kausalzusammenhänge zwar vorhanden, aber nur schwach ist. Die Vorhersagen sind in der unteren Hälfte von Abbildung 3.24 zu sehen.

Nach ihren Prüfurteilen wurden die Probanden zusätzlich gebeten, die Zusammenhänge zwischen den Ereignissen einzuschätzen. Sofern sie ihr Urteil auf die Prüfung der vermuteten Zusammenhänge stützen, sollte sich die Beurteilung der Hypothesen über die geschätzten Kontingenzen zwischen den kausal verbundenen Ereignissen vorhersagen lassen. Ist eine der beiden vermuteten Kausalrelationen nicht vorhanden oder nur schwach, wie dies bei den Hypothesen mit der Kennziffer 2 der Fall ist, dann sollte das Urteil niedrig ausfallen. Sind beide Zusammenhänge stark, dann sollte ein hohes Urteil resultieren. Der Zusammenhang zwischen den beiden nicht kausal verbundenen Ereignissen sollte sich als nicht relevant erweisen.

Prüfen die Probanden aber strukturelle Implikationen, dann sollte sich das Hypothesenprüfurteil über die Stärke des Zusammenhangs zwischen den nicht direkt kausal verbundenen Ereignissen vorhersagen lassen. Liegt bei den Gemeinsamer-Effekt Hypothesen ein Zusammenhang zwischen den Ursachen vor, dann sollten diese Hypothesen abgelehnt werden. Dagegen sollte bei einer Gemeinsame-Ursache Hypothese ein Zusammenhang zwischen den Effekten zur Annahme der Hypothese führen.

Probanden und Design

80 Studierende aller Fachbereiche nahmen an dem Experiment teil. Sie wurden über Aushänge angeworben und erhielten DM 10,- für ihre Teilnahme. Psychologiestudenten bekamen wahlweise eine Versuchspersonenstunde testiert. Keiner der Probanden hatte zuvor an einem der anderen Experimente teilgenommen. Die Probanden wurden per Zufall einer der acht Versuchsbedingungen zugewiesen und einzeln untersucht.

Versuchsablauf und Material

Der Ablauf und die Instruktionen wurden aus Experiment 4 übernommen. Alle Probanden bekamen die Aufgabe gestellt, zu überprüfen, ob die ihnen vorgelegte kausale Strukturhypothese zutrifft oder nicht. Nach der schriftlichen Instruktion wurde wiederum das Instruktionsverständnis überprüft und die Probanden angehalten, sich die Hypothese aufzuzeichnen. Danach erhielten sie 100 Karten mit Abbildungen der Aktivitäten der einzelnen Kerne zu dem jeweiligen Messzeitpunkt. Erneut wurde ihnen erlaubt, die Karten zu ordnen und sich Notizen zu machen. Nach ihrer Analyse der Daten wurden sie gebeten, ihre Beurteilung der Hypothese durch Ankreuzen auf einer Skala von 0 (trifft nicht zu [0%] zu) bis 100 (trifft sicher [zu 100%] zu) anzugeben. Zu ihrem Urteil erhielten die Probanden keine Rückmeldung. Danach wurden sie gebeten, Fragen zu den Zusammenhängen zwischen den Ereignissen zu beantworten. Die Fragen wurden paarweise vorgegeben. Alle Fragen hatten folgende Form: „Stell dir vor, dass der Kern X 100 mal aktiv ist. In wie vielen dieser Fälle ist auch Kern Y aktiv? Stell Dir vor, dass der Kern X in 100 Fällen nicht aktiv ist. In wie vielen dieser Fälle ist Kern Y aktiv?“⁴⁶ Die beiden vermuteten Kausalzusammenhänge wurden jeweils in Richtung des Kausalzusammenhangs abgefragt. Das heißt, in der Gemeinsame-Ursache Bedingung wurden die Probanden z.B. nach der Häufigkeit der Aktivität des Nukleus anterior bei Aktivität und Inaktivität des Pulvinars gefragt. In der Gemeinsamer-Effekt Bedingung wurden sie in der Gegenrichtung befragt: „Wie häufig ist der Pulvinar aktiv, wenn der Nukleus anterior 100 mal aktiv (inaktiv) ist?“ Dieser Wechsel der Fragerichtung ist deshalb angezeigt, weil anzunehmen ist, dass die Probanden die Kausalzusammenhänge nur vorwärts von der Ursache zur Wirkung hin repräsentieren (vgl. Waldmann & Martignon, 1998). Für den Zusammenhang zwischen den beiden nicht kausal verbundenen Ereignissen wurde dieselbe Art von Fragen gestellt. Dabei wurde die Fragerichtung ausbalanciert.

⁴⁶ Diese Form der Abfrage soll das Wissen über Kausalzusammenhänge am besten erfassen (Gigerenzer & Hoffrage, 1995; Cheng, persönliche Mitteilung).

Präsentierte Daten

Die Daten wurden über zwei Bayesianische Netzwerke generiert. Diese beruhen auf der Gemeinsame-Ursache und der Gemeinsamer-Effekt Struktur 1. Die Basisraten der Ursachen wurden auf $P(u) = 0,5$, die bedingte Wahrscheinlichkeit des Effektes bei Vorliegen einer Ursache auf $P(e|u) = P(e|u_1.\sim u_2) = P(e|\sim u_1.u_2) = 0,60$ und bei Abwesenheit der Ursachen auf $P(e|\sim u) = P(e|\sim u_1.\sim u_2) = 0,00$ festgelegt. Das Vorliegen mindestens einer Ursache war also bei beiden Datensätzen notwendig für das Auftreten der Effekte. Mit diesen bedingten Wahrscheinlichkeiten ergibt sich wie im Vorgängerexperiment in den Daten ‚Pro Gemeinsame-Ursache‘ eine unbedingte Kontingenz von $\Delta P = 0,60$ zwischen der Ursache und ihren beiden Effekten und in den Daten ‚Pro Gemeinsamer-Effekt‘ eine bedingte Kontingenz von $\Delta P = 0,60$ zwischen jeder der beiden Ursachen und dem Effekt. Abbildung 3.25 fasst die verwendeten Werte zusammen.

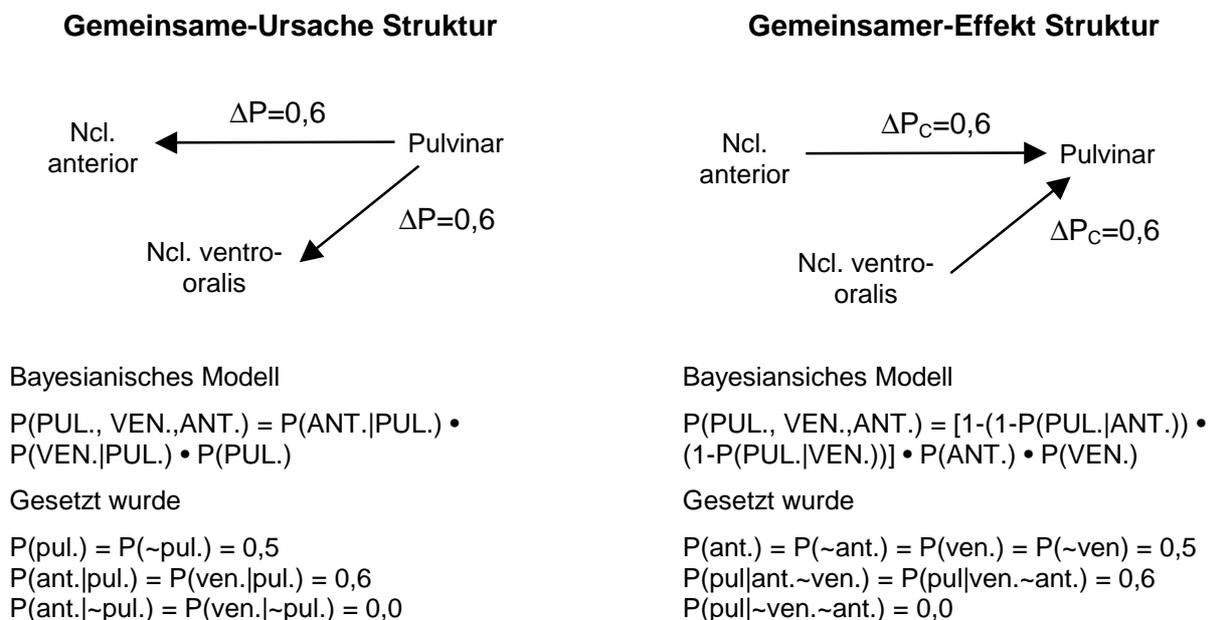


Abb. 3.25: Konstruktionsprinzipien der Daten für Experiment 5. Die Bayesianischen Modelle beziehen sich auf Ereignisvariablen (Großbuchstaben), die gesetzten Wahrscheinlichkeiten auf Ereignisausprägungen (Kleinbuchstaben)

Auf Grundlage dieser Modelle wurde je ein Datensatz mit 100 Fällen generiert. Abbildung 3.26 zeigt die Verteilung der Ereigniskombinationen.

Daten Pro Gemeinsame-Ursache					Daten Pro Gemeinsamer-Effekt				
	ant.		~ant.			ant.		~ant.	
	ven.	~ven.	ven.	~ven.		ven.	~ven.	ven.	~ven.
pul	18	12	12	8	pul	21	15	15	0
~pul.	0	0	0	50	~pul.	4	10	10	25

Abb. 3.26: Daten zu Experiment 5

Die Notwendigkeitsbeziehung ist an den Fällen zu erkennen, die bei den jeweiligen Daten nicht vorkommen. Bei den Daten ‚Pro Gemeinsame-Ursache‘ tritt nie eine Aktivität in den beiden Nuklei auf, wenn der Pulvinar inaktiv ist. Genauso ist bei den Daten ‚Pro Gemeinsamer-Effekt‘ der Pulvinar nie aktiv, wenn die beiden Nuklei inaktiv sind. Für die Zusammenhänge in den Daten ergibt sich folgendes Bild.

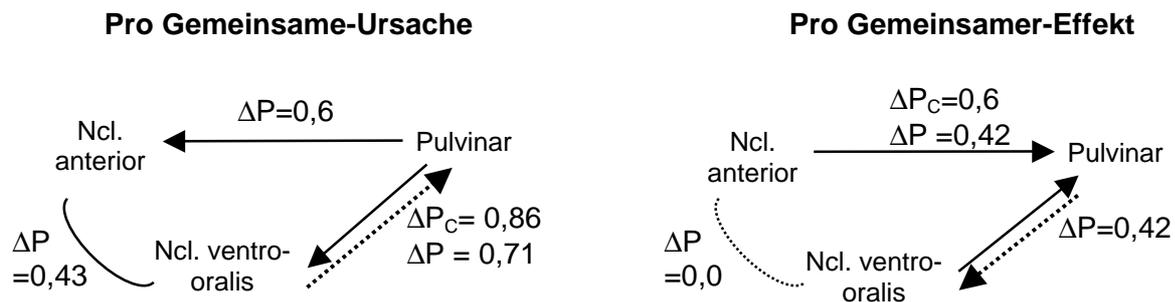


Abb. 3.27: Zusammenhänge zwischen den Ereignissen in Experiment 5

Für die Frage nach der strukturellen Sensitivität ist ja vor allem der Unterschied in den Zusammenhängen zwischen den nicht direkt kausal verbundenen Ereignissen wichtig. Dieser beträgt beim Gemeinsame-Ursache Modell $\Delta P = 0,43$ und beim Gemeinsamer-Effekt Modell $\Delta P = 0,0$.

Ergebnisse

Abbildung 3.28 gibt die Mittelwerte der Prüfurteile für alle Versuchsbedingungen an.

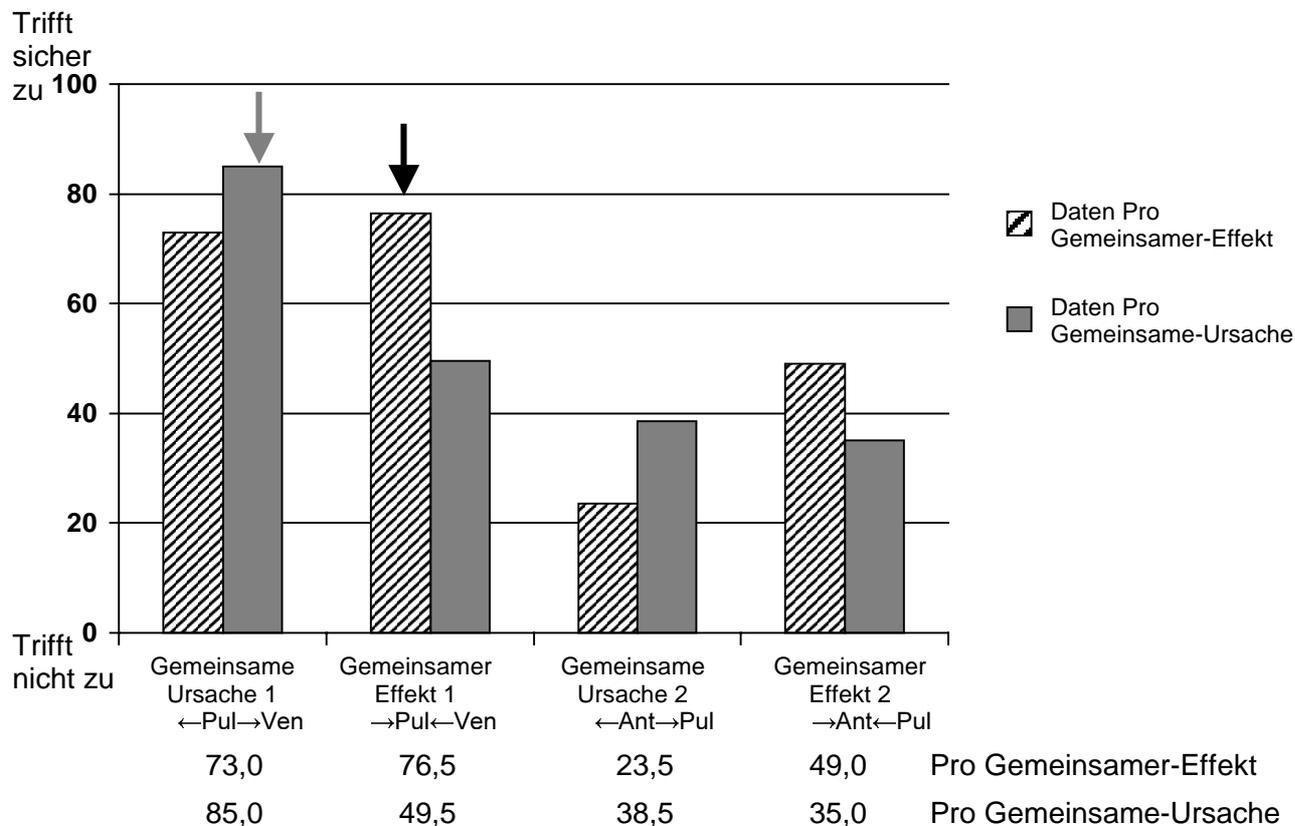


Abbildung 3.28: Ergebnisse Experiment 5. Mittelwerte der Prüfurteile.
Die Pfeile kennzeichnen das Urteil für die normativ korrekte Strukturhypothese

Sprechen die Daten für die Gemeinsamer-Effekt Hypothese 1, so beurteilen die Versuchspersonen die Gemeinsame-Ursache und die Gemeinsamer-Effekt Hypothese 1 zustimmend, während sie die beiden Hypothesen mit den Kennziffer 2 ablehnen. Ein anderes Bild zeigt sich für die Daten ‚Pro Gemeinsame-Ursache‘. Hier wird nur die Gemeinsame-Ursache Hypothese 1 angenommen, alle anderen werden zurückgewiesen.

Eine Varianzanalyse mit den Zwischensubjektfaktoren Strukturhypothese und Datensatz erbrachte einen signifikanten Haupteffekt für das vermutete Strukturmodell $F(3,72) = 17,92, p < 0,01$, und eine signifikante Interaktion zwischen gezeigten Daten und dem zu überprüfenden Modell, $F(3,72) = 4,05, p < 0,01, MSE = 512,9$. Zwischen den beiden Datenbedingungen ergab sich kein Unterschied ($F < 1$).⁴⁷ Diese Ergebnisse sprechen dafür, dass die Probanden zwischen den Strukturhypothesen unterschieden und die Daten je nach Hypothese unterschiedlich verrechneten. Betrachtet man die Verteilung der Mittelwerte, so

⁴⁷ Dieser Befund stützt die Vermutung, dass der Effekt der Datensätze in Experiment 4 auf Zufall beruht.

fällt auf, dass bei beiden Datensätzen die nicht zutreffenden Hypothesen mit der Kennziffer 2 richtigerweise abgelehnt wurden. Die Wahrscheinlichkeit ihres Zutreffens wurde auf unter 50% geschätzt. Für die Hypothesen mit der Kennziffer 1 zeigte sich ein differenzierteres Bild. Diese beiden Hypothesen stimmten darin überein, dass sie Kausalzusammenhänge zwischen denselben Arealen vermuteten und diese durch beide Datensätze bestätigt wurden. Eine Analyse nur für diese beiden letztgenannten Hypothesen erbrachte eine signifikante Interaktion zwischen Daten und Hypothese, $F(1,36) = 7,14$, $p < 0,05$, und einen Haupteffekt für die gezeigte Hypothese, $F(1,36) = 4,80$, $p < 0,05$, $MSE = 532,9$. Auffällig ist insbesondere die Form der Interaktion. Während die Probanden bei den Daten ‚Pro Gemeinsame-Ursache‘ die zugehörige Strukturhypothese als einzige annahmen, machten sie bei den Daten ‚Pro Gemeinsamer Effekt‘ keinen Unterschied zwischen der Gemeinsame-Ursache und Gemeinsamer-Effekt Hypothese 1. Das Ergebnis bei den Daten ‚Pro Gemeinsame-Ursache‘ spricht für eine Berücksichtigung struktureller Implikationen oder logischer Beziehungen. Die Gemeinsamer-Effekt Hypothese wird abgelehnt, obwohl beide vermuteten Kausalzusammenhänge vorlagen und eine erhebliche Stärke aufwiesen ($\Delta P_C = 0,86!$). Das Ergebnis für die Daten Pro Gemeinsamer-Effekt deutet in eine andere Richtung. Obwohl eine Aktivität in den beiden Nuklei notwendig für die Aktivität des Pulvinars war, machten die Probanden keinen Unterschied zwischen den beiden Hypothesen mit der Kennziffer 1. Dieser Befund spricht sowohl gegen eine Berücksichtigung struktureller Implikationen des jeweiligen Modells als auch gegen eine Nutzung logischer Kriterien.

Wie beurteilten die Probanden die Zusammenhänge zwischen den Ereignissen? Die Einschätzungen der relativen Häufigkeiten des Effektes in An- bzw. Abwesenheit der jeweiligen Ursache wurden hierzu in Kontingenzen umgerechnet. Abbildung 3.29 gibt einen Überblick über die Befunde.

	Kontingenz Pulvinar - Anterior	Kontingenz Ventreoralis - Pulvinar	Kontingenz Anterior - Ventreoralis
Daten Pro Gemeinsamer-Effekt	29,9	21,8	1,88
Daten Pro Gemeinsame-Ursache	34,5	37,5	5,88

*Abb. 3.29: Geschätzte Kontingenzen zwischen den Aktivitäten der Thalamuskern
(Werte in Prozent: $\Delta P \times 100$)*

Eine Varianzanalyse mit Daten als Zwischensubjektfaktor und Zusammenhang als Innersubjektfaktor ergab einen signifikanten Effekt für die Zusammenhänge, $F(2,156) = 52,7$,

$p < 0,01$, $MSE = 371,9$, und einen signifikanten Haupteffekt für die gezeigten Datensätze, $F(1,78) = 4,42$, $p < 0,05$, $MSE = 895,5$. Die Interaktion wurde nicht signifikant, $F(2,156) = 2,35$, $p = 0,10$, $MSE = 371,9$. Der Effekt für die beiden Datensätze war zu erwarten gewesen. Die Fragen bezogen sich jeweils auf die unbedingten relativen Häufigkeiten. Die unbedingten Kontingenzen für die Zusammenhänge zwischen dem Pulvinar und den beiden anderen Kernen betragen bei den Daten ‚Pro Gemeinsamer-Effekt‘ $\Delta P = 0,42$ und bei den Daten ‚Pro Gemeinsame-Ursache‘ $\Delta P = 0,60$. Es liegt also bereits in den Daten ein Unterschied vor. Daneben erkennen die Probanden auch richtig, dass der Zusammenhang zwischen den beiden Nuklei schwächer ist als zwischen dem Pulvinar und den beiden Nuklei. Wichtig für die Prüfung der strukturellen Implikationen ist das Vorliegen bzw. Fehlen eines Zusammenhangs zwischen dem Nukleus anterior und dem Nukleus ventrooralis. Ein geplanter Vergleich dieser beiden Zellen zeigte keinen signifikanten Unterschied. Die Probanden halten offenbar diese beiden Ereignisse in beiden Datenbedingungen für unabhängig.

Lassen sich die Prüfurteile über die Einschätzungen der Kausalzusammenhänge vorhersagen? Zur Beantwortung dieser Frage wurden zwei Regressionsanalysen durchgeführt, eine für die beiden Gemeinsame-Ursache und eine für die beiden Gemeinsamer-Effekt Modelle. Als Prädiktoren wurden die geschätzte Stärke der beiden vermuteten Kausalzusammenhänge und die Einschätzung des Zusammenhangs zwischen den beiden laut Hypothese nicht direkt kausal verbundenen Ereignissen verwendet.⁴⁸ Das Kriterium war die jeweilige Beurteilung der vorgelegten Strukturhypothese. Abbildung 3.30 stellt die Ergebnisse dar.

⁴⁸ Es wurden nicht die Einschätzungen der drei Zusammenhänge per se verwendet, da diese je nach Modell eine andere Bedeutung haben. So stand der Zusammenhang von Pulvinar und Nucleus ventrooralis bei der Gemeinsame-Ursache Hypothese 1 für einen vermuteten Zusammenhang, bei der Gemeinsame-Ursache Hypothese 2 aber für die implizierte Kovariation zwischen den Effekten. Durch die Verwendung der beiden vermuteten und des nicht vermuteten Zusammenhangs als Prädiktoren wurde diesem Unterschied Rechnung getragen.

Gemeinsame-Ursache Hypothese		Gemeinsamer-Effekt Hypothese	
Varianzaufklärung	$R^2 = 0,43$	Varianzaufklärung	$R^2 = 0,30$
Signifikanztest	$F(3,39) = 9,21$ $p < 0,01$ $MSE = 580,6$	Signifikanztest	$F(3,39) = 5,05$ $p < 0,01$ $MSE = 662,7$
Vermuteter Kausalzusammenhang 1	$\beta = 0,23$	Vermuteter Kausalzusammenhang 1	$\beta = 0,32$
Vermuteter Kausalzusammenhang 2	$\beta = 0,45^*$	Vermuteter Kausalzusammenhang 2	$\beta = 0,26$
Zusammenhang nicht kausal verbundene Ereignisse	$\beta = -0,13$	Zusammenhang nicht kausal verbundene Ereignisse	$\beta = -0,37^*$

Abb.3.30: Regressionsanalysen Experiment 5. Vorhersage des Prüfurteils über die geschätzten Kontingenzen zwischen den Ereignissen.

*Die mit * gekennzeichneten β -Werte sind auf dem 5% Niveau signifikant.*

Beide Analysen ergaben, dass sich das Prüfurteil über die Einschätzungen der Zusammenhänge vorhersagen ließ. Ein beträchtlicher Anteil an Varianz konnte darüber aufgeklärt werden, $R^2 = 0,43$ bei den Gemeinsame-Ursache Hypothesen und $R^2 = 0,30$ bei den Gemeinsamer-Effekt Hypothesen. Dieses Ergebnis spricht eindeutig dafür, dass Probanden die Stärke der vermuteten Kausalzusammenhänge als Kriterium zur Prüfung der Strukturhypothese heranzogen.

Die Regressionsgewichte zeigten ein uneinheitliches Bild. Bei den Gemeinsame-Ursache Hypothesen wies der zweite der vermuteten Kausalzusammenhänge ein signifikantes Beta-Gewicht auf. Dies war der Zusammenhang, der bei den Daten ‚Pro Gemeinsamer-Effekt‘ nicht bestätigt wurde. Offensichtlich haben die Probanden dies erkannt und lehnten deshalb die Hypothese ab. Für die Gemeinsamer-Effekt Hypothesen ergab sich ein anderes Bild. Während beide vermuteten Zusammenhänge knapp die Signifikanz verfehlten ($p < 0,10$), erwies sich der Zusammenhang der nicht kausal verbundenen Ereignisse als negativer Prädiktor.

Die Ergebnisse stimmen insgesamt in großen Teilen mit denen der vorherigen Experimente überein. Die Strukturhypothesen, bei denen einer der beiden vermuteten Kausalzusammenhänge entweder gar nicht, oder nur schwach vorhanden war, wurden abgelehnt. Waren beide erwarteten Kausalzusammenhänge vorhanden, so wurden in drei von 4 Fällen die Strukturhypothesen angenommen. Ein Teilurteil weicht aber von den Vorgänger-experimenten erheblich ab. Bei den Daten ‚Pro Gemeinsame-Ursache‘ wurde nur

die normativ richtige Gemeinsame Ursache-Strukturhypothese 1 angenommen. Alle anderen Hypothesen, auch die Gemeinsamer-Effekt Hypothese 1, bei der beide Kausalzusammenhänge vorlagen, wurden zurückgewiesen.

Wie ist dieser Befund zu erklären? Ein erster Erklärungsversuch könnte die Berücksichtigung der strukturellen Implikationen sein. Diese sind nur für die bevorzugte Hypothese erfüllt. Gegen diese Erklärung sprechen allerdings mehrere Gründe. Erstens nahmen die Probanden bei beiden Datensätzen die Gemeinsame-Ursache Hypothese 1 an. Dagegen wurde die Gemeinsamer-Effekt Hypothese bei den Daten ‚Pro Gemeinsamer-Effekt‘ angenommen und bei den Daten ‚Pro Gemeinsame-Ursache‘ abgelehnt. Weshalb aber sollte die Sensitivität für strukturelle Implikationen nur auf Gemeinsamer-Effekt Hypothesen beschränkt sein? Ebenfalls gegen eine Sensitivität spricht, dass die Probanden den Zusammenhang zwischen dem Nukleus anterior und dem Nukleus ventrooralis in beiden Datenbedingungen gleich beurteilten. Voraussetzung für eine Prüfung anhand struktureller Implikationen ist es aber, die implizierte Abhängigkeit bzw. Unabhängigkeit genau zu überprüfen. Wenn die Probanden dies gemacht hätten, so hätten sie einen Unterschied zwischen den beiden Datensätzen finden müssen.⁴⁹ Ein drittes Argument gegen das Vorliegen einer strukturellen Sensitivität liefern die Regressionsanalysen. Wenn die Probanden sensitiv wären, dann sollte sich bei den Gemeinsame-Ursache Hypothesen ein positives und bei den Gemeinsamer-Effekt Hypothesen ein negatives Regressionsgewicht für den nicht vermuteten, aber implizierten Zusammenhang ergeben. Diese Vorhersage konnte nur für die Gemeinsamer-Effekt Hypothesen bestätigt werden. Allerdings zeigte sich bei beiden Hypothesentypen ein negatives Regressionsgewicht, was auch mit der Interpretation verträglich ist, dass das Vorliegen nicht vermuteter Zusammenhänge generell eher zu einer Reduktion der Prüfurteile führt. Kurz gesagt, für eine systematische Nutzung struktureller Implikationen gibt es keine konvergierenden Evidenzen.

Eine zweite Erklärung für diesen besonderen Befund könnte in der logischen Struktur der Datensätze liegen. Dies erscheint vor allem deshalb plausibel, weil der Befund erstmals in diesem Experiment auftrat. Nutzen die Probanden die logische Notwendigkeitsbeziehung zwischen den Ereignissen als Kriterium für ihr Urteil? In den Daten ‚Pro Gemeinsame-Ursache‘ war die Aktivität des Pulvinars notwendig für eine Aktivität in den beiden anderen Nuklei. Bei diesen Daten wurde in der Tat nur die Gemeinsame-Ursache Hypothese 1 angenommen, nach der die eine Ursache notwendig für ihre beiden Effekte war. Allerdings nutzten die Probanden die logische Notwendigkeitsrelation nicht bei den Daten ‚Pro Gemeinsamer-Effekt‘. Obwohl hier die beiden Nuklei als Ursachen notwendig für eine

⁴⁹ Die Ergebnisse für die Zusammenhangsschätzungen deuten darauf hin, dass sich die Probanden nur einen groben Eindruck über die in den Daten gegebenen Zusammenhänge verschafft haben. Diese Interpretation würde auch die deutliche Unterschätzung der tatsächlichen Kontingenzen sowie deren erhebliche Streuung erklären.

Aktivität des Pulvinar waren, wurde die entsprechende Hypothese (Gemeinsamer-Effekt Hypothese 1) nicht der Gemeinsame-Ursache Hypothese 1 vorgezogen.⁵⁰ Es gibt also keinen Hinweis darauf, dass die Probanden bei den Daten ‚Pro Gemeinsamer-Effekt‘ die logischen Relationen als Entscheidungskriterium herangezogen haben. Aber weshalb sollten die Probanden nur bei den Daten Pro Gemeinsame-Ursache auf die logischen Relationen geachtet haben? Sie konnten ja nicht im Vorhinein wissen, auf welche Daten sie stoßen würden.

Dennoch liegt der Grund für die unterschiedlichen Urteile bei den beiden Datensätzen wahrscheinlich in den logischen Beziehungen. Diese wiesen je nach Datensatz eine unterschiedliche Komplexität auf. Während bei den Daten ‚Pro Gemeinsame-Ursache‘ das Vorliegen genau eines Ereignisses notwendig für die beiden anderen Ereignisse war, war bei den Daten ‚Pro Gemeinsamer-Effekt‘ das Vorliegen eines von zwei Ereignissen notwendig für das Auftreten des dritten. So wurde zum Beispiel der Pulvinar auch aktiv, wenn der Nukleus anterior nicht aktiv war, nämlich genau dann, wenn der Nukleus ventrooralis aktiv war. Damit ist die Notwendigkeitsrelation bei den Daten ‚Pro Gemeinsamer-Effekt‘ deutlich komplexer: Das Auftreten irgendeines von zwei Ereignissen ist notwendig, nicht das Auftreten eines bestimmten. Dieser Unterschied manifestiert sich auch deutlich in den Daten. Während bei den Daten ‚Pro Gemeinsame-Ursache‘ gleich drei mögliche Ereigniskombinationen nicht vorkommen, ist dies bei den Daten ‚Pro Gemeinsamer-Effekt‘ lediglich eine Ereigniskombination (Anwesenheit des Effektes bei Abwesenheit beider Ursachen). Dieser Unterschied könnte die Asymmetrie in den Prüfurteilen erklären. Wahrscheinlich bemerkten die Probanden bei den Daten ‚Pro Gemeinsame-Ursache‘ die dort vorhandene Notwendigkeit des einen spezifischen Ereignisses für die beiden anderen und nutzten diese Relation für ihre Entscheidung. Bei den Daten ‚Pro Gemeinsamer-Effekt‘ lag kein einzelnes notwendiges Ereignis vor. Dass eines von zwei Ereignissen notwendig war, wurde möglicherweise nicht erkannt oder zumindest nicht als Kriterium herangezogen.

Zusammenfassend gesagt: Auch die extreme Vereinfachung der präsentierten Daten führte nur teilweise zu einer Verbesserung der Prüfleistung. Nur wenn ein einziges bestimmtes Ereignis notwendig für das Auftreten anderer Ereignisse war, wurde diese logische Relation als Entscheidungskriterium herangezogen. Der Befund, dass Personen Strukturhypothesen nicht anhand ihrer strukturellen Implikationen testen, konnte in diesem Experiment erneut bestätigt werden.

⁵⁰ Einen weiteren Beleg dafür, dass die Probanden nicht generell auf die Notwendigkeitsrelation achten, liefern die ersten beiden Experimente. Bei diesen war im Datensatz ‚Pro Gemeinsamer-Effekt‘ ebenfalls eine Notwendigkeit der beiden Ursachen für den Effekt gegeben. Die Probanden haben dies aber in keinem dieser Experimente berücksichtigt.

Diskussion

Die Fragestellung für die Experimente 2 bis 5 war, ob Probanden zur Prüfung von Strukturhypothesen deren strukturelle Implikationen als Entscheidungskriterium nutzen. Strukturelle Implikationen sind, normativ gesehen, das einzige wirklich valide Kriterium. Zur Untersuchung dieser Frage wurden Probanden gebeten, verschiedene Strukturhypothesen anhand eines Datensatzes zu prüfen. Als Strukturhypothesen wurden das Gemeinsame-Ursache und das Gemeinsamer-Effekt Modell verwendet. Die präsentierten Daten waren so konstruiert, dass sie bei einem Teil der Hypothesen dem Vorliegen eines der vermuteten Kausalzusammenhänge widersprachen, bei einem anderen Teil die strukturellen Implikationen verletzten und jeweils eine der Hypothesen vollständig bestätigten. Der übereinstimmende Befund aller vier Experimente war, dass die Probanden strukturelle Implikationen nicht berücksichtigten. Sie ignorierten dieses entscheidende Kriterium. Dabei war es vollkommen unerheblich, ob sie mehrere Hypothesen nacheinander am selben Datensatz testen sollten (Experiment 2), zwischen zwei Hypothesen in direktem Vergleich entscheiden mussten (Experiment 3) oder nur eine einzige Hypothese zur Prüfung hatten (Experimente 4 und 5). Ebenfalls keinen Einfluss hatte es, ob in den Datensätzen von den strukturellen Implikationen nur die unbedingten Abhängigkeiten, diese aber in besonders salienter Weise, erfüllt waren (Experimente 2 und 3) oder aber alle Implikationen vollständig normativ erfüllt waren (Experiment 4). Selbst das zusätzliche Vorliegen logischer Notwendigkeitsbeziehungen führte nur im einfachsten Fall zu einer Verbesserung der Prüfleistung (Experiment 5).

Bisherige Forschungsarbeiten haben sich meines Wissens nicht mit der Frage des Prüfens von Strukturhypothesen beschäftigt. Die Untersuchungen haben sich wie in Kapitel 2.3 dargestellt auf das Prüfen einer oder mehrerer kausaler Zusammenhangshypothesen beschränkt. Die Ergebnisse dieser Studien haben gezeigt, dass Probanden (Kinder wie Erwachsene) durchaus in der Lage sind, mehrere Kausalhypothesen gleichzeitig zu testen.⁵¹ Die Ergebnisse der hier vorgelegten Experimente stimmen mit diesen Befunden überein und erweitern diese auf kausale Strukturhypothesen. Die Ergebnisse aller vier Experimente deuten darauf hin, dass die Probanden Strukturhypothesen testen, in dem sie überprüfen, ob die vermuteten direkten Kausalzusammenhänge vorliegen oder nicht. Die Prüfurteile in den Experimenten 2, 3, 4 lassen sich über die Anwendung dieser Strategie am besten erklären. Zusätzliche Evidenz lieferte die Regressionsanalyse von Experiment 5. Mehr als ein Drittel

⁵¹ Eine Vorbedingung hierfür war, dass der Sachverhalt neutral war und die Probanden kein spezifisches Vorwissen hatten. Dies war bei allen hier verwendeten Sachverhalten der Fall. Eine weitere Bedingung war, dass die Anforderungen an die Verarbeitungskapazität nicht zu hoch sein durften. Diesem Umstand wurde dadurch Rechnung getragen, dass die Probanden in allen Experimenten sich Notizen machen durften und die Bearbeitungszeit nicht begrenzt wurde.

der Varianz in den Urteilen kann über die geschätzte Stärke der vermuteten Kausalzusammenhänge aufgeklärt werden. Der Zusammenhang der nicht direkt kausal verbundenen Ereignisse scheint dagegen eine untergeordnete Rolle zu spielen. Das bedeutet, dass die Probanden wahrscheinlich Strukturhypothesen genau so testen, wie sie mehrere einzelne Zusammenhangshypothesen testen, nämlich über die Bestimmung der Kausalzusammenhänge. Dieses Vorgehen ist sowohl bei einzelnen Kausalzusammenhängen als auch bei mehreren Kausalrelationen normativ richtig, sofern das angenommene Modell zutrifft. Es ist aber für die normativ korrekte Prüfung der Gültigkeit von Strukturhypothesen ungeeignet.

Was bedeutet dieser Befund für unsere Frage nach der Sensitivität für verschiedene kausale Strukturmodelle?

Die Befunde zeigen, dass Probanden Kausalmodelle anhand von Daten sinnvoll überprüfen können. Das dabei eingesetzte Vorgehen ist allerdings eine Heuristik, die nicht unter allen Bedingungen zu normativ richtigen Schlüssen führt. Durch das Überprüfen des Vorliegens der vermuteten Kausalzusammenhänge können aber zumindest Modelle ausgeschlossen werden, die von nicht vorhandenen Kausalrelationen ausgehen. Strukturelle Implikationen werden nicht beachtet, obwohl diese valide Entscheidungskriterien sind. Die Probanden kennen offenbar die Bedeutung dieser für das Testen von Strukturhypothesen nicht. Weder versuchen sie, diese aktiv zu überprüfen, noch fällt ihnen auf, wenn diese verletzt werden. Dass die Probanden die normativen Prüfkriterien der verschiedenen Kausalmodelle nicht kennen, heißt aber nicht, dass sie die Unterschiede zwischen den Modellen nicht kennen. Es ist durchaus möglich, dass die Probanden einen Unterschied zwischen den Modellen machen, bevor sie die Prüfung beginnen. Da aber das zur Überprüfung gewählte Vorgehen nicht sensitiv für die strukturellen Implikationen der Modelle ist, zeigen sich keine Unterschiede zwischen Gemeinsame-Ursache- und Gemeinsamer-Effekt Modellen in den Prüfurteilen. Um zu überprüfen, ob diese Überlegung zutrifft, wurde zusätzlich eine Befragungsstudie durchgeführt.

3.2.5 Befragungsstudie

Ziel der Befragungsstudie ist es, herauszufinden, welche Unterschiede Probanden zwischen Gemeinsame-Ursache und Gemeinsamer-Effekt Modellen beim Prüfen von kausalen Strukturhypothesen machen. Im Gegensatz zu den Experimenten 2-5 erhielten die Probanden in dieser Studie nicht die Aufgabe, Strukturhypothesen zu prüfen, sondern lediglich die Aufgabe, mögliche Beobachtungen daraufhin zu beurteilen, wie sehr diese für oder gegen die jeweilige Hypothese sprechen.

Wie in Kapitel 2.3 dargestellt, liegen zahlreiche übereinstimmende Befunde darüber vor, wie Probanden Evidenzen in Bezug auf einzelne Kausalzusammenhänge werten. So wird das gemeinsame Auftreten von Ursache und Effekt als positiver Beleg für das Vorhandensein eines Zusammenhangs gewertet. Negative Evidenzen sind das alleinige Auftreten der Ursache oder des Effektes. Das gemeinsame Ausbleiben beider Ereignisse wird in der Regel als neutrale oder leicht positive Evidenz gewertet. Es ist zu erwarten, dass die Probanden dieses Bewertungsschema auch auf kausale Strukturmodelle übertragen. Für eine Gemeinsame-Ursache Hypothese sprechen danach Fälle, in denen sowohl die Ursache als auch die beiden Effekte beobachtet werden. Alle anderen möglichen Beobachtungsfälle, in denen nur die Ursache oder nur die Effekte gegeben sind, sprechen gegen diese Hypothese. Für eine Gemeinsamer-Effekt Hypothese würden Fälle sprechen, in denen mindestens eine der beiden Ursachen und der Effekt gegeben sind. Fälle, in denen nur der Effekt oder nur die Ursachen vorliegen, sind dagegen negative Evidenzen. Beobachtungen, bei denen weder die Ursachen noch die Effekte vorliegen, sollten entweder neutral oder als positive Evidenzen für beide Strukturhypothesen gewertet werden. Damit unterscheiden sich die Vorhersagen der beiden Strukturhypothesen in den Fällen, in denen nur zwei der drei Ereignisse auftreten. Nehmen wir an, die Gemeinsame-Ursache Hypothese ist, dass das Ereignis C die Ereignisse A und B hervorruft. Die Gemeinsamer-Effekt Hypothese geht genau von entgegengesetzten Kausalzusammenhängen aus, die Ereignisse A und B sind die Ursachen von C. Eine mögliche Beobachtung ist es, dass nur die Ereignisse A und C vorliegen. Nach der Gemeinsame-Ursache Hypothese ist dies ein Fall, in dem die Ursache (C) nur von einem ihrer Effekte (A) gefolgt wird. Dies sollte aber gemäß der Hypothese nicht geschehen. Nach der Gemeinsamer-Effekt Hypothese ist dies ein Fall, in dem eine der beiden Ursachen (A) den Effekt (C) hervorgerufen hat. Eine solche Beobachtung würde für die Hypothese sprechen. Dieselbe Ereigniskombination ist also je nach Strukturhypothese entweder eine positive Evidenz oder eine negative. Ob Probanden auch diesen Unterschied machen, soll die Befragungsstudie zeigen.

Probanden

Sechzig Studierende aller Fachbereiche wurden einzeln befragt. Ihnen wurde per Zufall eine der beiden Strukturhypothesen vorgelegt. Für ihre Teilnahme erhielten sie wahlweise eine halbe Versuchspersonenstunde oder DM 6,--.

Durchführung und Material

Als Hintergrundgeschichte wurde der aus den Experimenten 2-5 bekannte Kausalzusammenhang zwischen den drei Thalamuskernen verwendet. Die Instruktion wurde ebenfalls übernommen. Die beiden Strukturhypothesen waren die Gemeinsame-Ursache Hypothese 1, nach der die Aktivität des Pulvinar die Ursache für die Aktivität in den beiden Nuklei ist, und die Gemeinsamer-Effekt Hypothese 1, nach der die Aktivität des Pulvinar der Effekt der Aktivitäten in den beiden Nuklei ist. Abweichend von der ursprünglichen Version wurde den Probanden mitgeteilt, dass ihnen nun eine Reihe von möglichen Beobachtungen gezeigt werde. Ihre Aufgabe sei es zu beurteilen, ob die jeweilige Beobachtung für oder gegen die Hypothese spreche. Daraufhin wurden ihnen in zufälliger Reihenfolge alle acht möglichen Kombinationen von Aktivitäten in den drei Arealen gezeigt. Dazu wurden die in den vorherigen Experimenten eingesetzten fingierten PET Scan Aufnahmen verwendet. Die Probanden sollten ihr jeweiliges Urteil auf einer Skala von -10 (spricht stark gegen die Hypothese, d.h., wenn die Hypothese zutrifft, dürfte dieser Fall nicht auftreten) bis +10 (spricht stark für die Hypothese, d.h., wenn die Hypothese zutrifft, sollte der Fall häufig auftreten) angeben. Die Null wurde als „spricht weder für noch gegen die Hypothese“ definiert. Die Probanden erhielten zu ihren Urteilen keinerlei Rückmeldung.

Ergebnisse

Abbildung 3.31 zeigt die mittleren Einschätzungen der Probanden. Waren alle drei Kerne aktiv, so werteten die Probanden dies als stark positive Evidenz für beide Hypothesen. Sie stimmten ebenso darin überein, dass die Aktivität nur in einem der Kerne als negative Evidenz zu werten ist. Das heißt, sie waren der Meinung, dass bei Zutreffen der Hypothese weder eine Ursache noch ein Effekt alleine auftreten sollten. Interessant ist der Fall, in dem nur die beiden Nuklei aktiv waren. Für die Gemeinsame-Ursache Hypothese bedeutete dies, dass beide Effekte ohne ihre Ursache auftraten, für die Gemeinsamer-Effekt Hypothese, dass der Effekt ausblieb, obwohl beide Ursachen da waren. Trotz dieser extrem unterschiedlichen Bedeutung wurde dieser Fall beide Mal als der Hypothese widersprechend angesehen. Wenn keine Aktivität beobachtet wurde, so wurde dies im Mittel als leicht positive Evidenz für beide Hypothesen angesehen. Hinter diesem Wert standen allerdings zwei unterschiedliche Ansichten. Ein Teil der Probanden sah dies als positiven Beleg, die anderen als neutral.

Beobachtete Aktivität			Gemeinsame-Ursache Hypothese	Gemeinsamer-Effekt Hypothese
Nukleus Anterior	Nukleus Ventro-oralis	Pulvinar		
✓	✓	✓	9,55	8,77
✓	✓	-	-6,00	-7,45
✓	-	✓	1,52*	6,06*
-	✓	✓	2,71*	6,71*
✓	-	-	-3,90	-5,32
-	✓	-	-4,19	-6,03
-	-	✓	-5,84	-4,58
-	-	-	3,32	2,90

*Abb. 3.31: Ergebnisse der Befragungsstudie: Mittelwerte der Beurteilung der einzelnen möglichen Beobachtungen. Positive Werte bedeuten, dass die Beobachtung für, negative, dass sie gegen die Hypothese spricht. Die mit * gekennzeichneten Urteile unterscheiden sich auf dem 5% Signifikanzniveau.*

Bei zwei Fällen unterschieden sich die Urteile für die beiden Hypothesen: Aktivität im Pulvinar und in nur einem der beiden Nuklei. Für die Gemeinsamer-Effekt Hypothese bedeutete dies, dass eine der beiden Ursachen und der Effekt vorhanden waren. Dies wurde als positive Evidenz gewertet. Für die Gemeinsame-Ursache Hypothese bedeutete dieser Fall, dass die Ursache und nur einer der beiden Effekte vorhanden waren. Dies wurde als weder positive noch negative Evidenz gewertet. Eine Folge von univariaten Varianzanalysen erbrachte nur für diese beiden Fälle einen signifikanten Unterschied, $F_{\text{Ncl.ant.-Pul.}}(1,59) = 7,16$, $p < 0,05$, $MSE = 44,8$; $F_{\text{Ncl.ven.-Pul.}}(1,59) = 6,31$, $p < 0,05$, $MSE = 39,3$.

Diese Ergebnisse unterstützen die Vermutung, dass die Probanden mögliche Evidenzen bei kausalen Strukturhypothesen ähnlich wie bei kausalen Zusammenhangshypothesen bewerten. Das Gemeinsame Auftreten von Ursachen und Effekten ist eine positive Evidenz, das alleinige Auftreten nur eines von beiden ist eine negative Evidenz. Auch eine im Schnitt leicht positive Bewertung des gemeinsamen Ausbleibens wurde gefunden. Ebenso bestätigt wird die Vorhersage, dass sich bei diesem Bewertungsschema Unterschiede zwischen den verschiedenen Strukturhypothesen zeigen müssten.

Diskussion

Die Probanden hatten die Aufgabe für eine Strukturhypothese zu bestimmen, welche der möglichen Beobachtungsfälle für oder gegen die Hypothese sprechen. Es wurde dabei eine Gemeinsame-Ursache mit einer Gemeinsamer-Effekt Hypothese kontrastiert. Die Ergebnisse sprechen dafür, dass Probanden zwischen diesen einen Unterschied machten. Während bei einer Gemeinsamer-Effekt Hypothese das Auftreten von nur einer der beiden Ursachen und des Effektes als Evidenz für die Hypothese gewertet wurde, beurteilten die Probanden das Auftreten der einen Ursache und nur eines der beiden Effekte als neutral für eine Gemeinsame-Ursache Hypothese. Dabei bezogen sich beide Gruppen von Probanden mit ihren Urteilen auf dieselben Fälle.

Diese Befunde erweitern die bisher vorliegenden Erkenntnisse für die Bewertung von Evidenzen für kausale Zusammenhangshypothesen. Dabei zeigte sich eine weitgehende Übereinstimmung. Das gemeinsame Auftreten der beiden Ereignisse wird als stark positive Evidenz gewertet, das alleinige Auftreten nur eines der Ereignisse wird negativ bewertet und eine gemeinsame Abwesenheit der Ereignisse wird als leicht positive Evidenz angesehen.

Die Urteile sind daher auch mit These vereinbar, dass die Probanden kausale Strukturmodelle als die Summe einzelner Kausalzusammenhänge ansehen. Nehmen wir das Gemeinsame-Ursache Modell und den Fall, dass nur die Ursache und einer der beiden Effekte gegeben sind. Dieser Fall spricht für den einen der beiden Kausalzusammenhänge, er spricht aber gleichzeitig gegen den anderen Kausalzusammenhang. Die insgesamt neutrale Beurteilung könnte eine Folge dieser beiden einander entgegengesetzten Evidenzen sein. Bei einem Gemeinsamer-Effekt Modell spricht der Fall, dass nur eine Ursache und der Effekt vorliegen, eindeutig für den einen der beiden Kausalzusammenhänge. Über die Wirkung der nicht vorhandenen Ursache kann dabei keine Aussage gemacht werden. Im Gegensatz zum Gemeinsame-Ursache Modell liegt bei dieser Beobachtung kein Konflikt zwischen den Evidenzen für die beiden Kausalzusammenhänge vor. Daher sollten Probanden diese Evidenz uneingeschränkt positiv werten, was tatsächlich der Fall war. Diese Analyse zeigt, dass die Probanden, auch wenn sie das Modell als Summe einzelner Kausalzusammenhänge auffassen, dennoch mögliche Beobachtungen modellspezifisch, normativ richtig beurteilen können.

Was bedeutet dieser Befund für unsere Frage nach der Sensitivität für verschiedene Kausalmodelle?

Die Frage, die sich nach den Experimenten 2 bis 5 stellte, war, ob die Probanden einen Unterschied zwischen einer Gemeinsamen-Ursache und einer Gemeinsamer Effekt Hypothese machen. Diese Frage lässt sich nun aufgrund der Ergebnisse der Befragungsstudie klar bejahen. Probanden unterscheiden die Modelle anhand ihrer Struktur,

auch wenn es um das Testen von Hypothesen geht. Sie beurteilen korrekt, ob eine mögliche Beobachtung für die jeweilige Strukturhypothese eine positive oder negative Evidenz ist.⁵²

Die unterschiedlichen Beurteilungen der möglichen Beobachtungen deuten zusätzlich darauf hin, dass mit der jeweiligen Strukturhypothese spezifische Erwartungen verbunden werden. Ist ein Gemeinsamer-Effekt Modell gegeben, so scheinen die Probanden Fälle zu erwarten, in denen beide Ursachen und der Effekt, eine der Ursachen und der Effekt oder weder die Ursachen noch der Effekt zu beobachten sind. Das sind auch normativ gesehen genau die Fälle, die bei einer solchen Struktur vor allem auftreten sollten. Ist ein Gemeinsame-Ursache Modell gegeben, dann scheinen die Probanden Fälle zu erwarten, bei denen die Ursache und beide Effekte oder weder die Ursache noch die Effekte gegeben sind. Fälle, in denen die Ursache und nur ein Effekt auftritt, sollten nicht beobachtet werden. Diese Erwartungen würden sehr gut mit den strukturellen Implikationen der Modelle übereinstimmen. Bei einem Gemeinsamer-Effekt Modell werden Fälle erwartet, bei denen beide, eine oder keine Ursache gegeben sind. Dies bedeutet, dass die Probanden keine spezifischen Vorannahmen bezüglich einer Kovariation der Ursachen haben. Bei einem Gemeinsame-Ursache Strukturmodell erwarten sie vor allem Fälle, bei denen entweder beide Effekte oder keiner der beiden Effekte gegeben ist. Dies ist aber gleichbedeutend mit der Erwartung einer Kovariation der Effekte. Die Beurteilung der möglichen Beobachtungen deutet also darauf hin, dass Personen zumindest bei ihren Erwartungen eine gewisse Sensitivität für strukturelle Implikationen aufweisen. Ziel der folgenden Experimente ist es, weitere und direktere Belege für diese strukturelle Sensitivität zu finden.

⁵² Dies widerlegt aber nicht die Befunde der vorangegangenen Experimente. Die Ergebnisse der Studie zeigen, dass Probanden einen Unterschied zwischen den Modellen machen und die Ergebnisse der Experimente zeigen, dass diese Unterschiede bei der Prüfung von kausalen Strukturhypothesen nicht zum Tragen kommen.

3.3 Experimente zur Vorhersage strukturell implizierter Zusammenhänge

Der Befund der Experimente 2-5 war, dass Probanden beim Prüfen kausaler Strukturhypothesen nicht sensitiv für die Verletzung struktureller Implikationen sind. Die Frage, die sich an diese Untersuchungen anschließt ist, ob Probanden bei anderen Aufgaben sensitiv für strukturelle Implikationen von Kausalmodellen sind. Eine solche Aufgabe ist es, separat gelernte Kausalzusammenhänge in ein einheitliches Modell zu integrieren und aus diesem Modell Vorhersagen über den implizierten Zusammenhang zwischen den nicht direkt kausal verbundenen Ereignissen abzuleiten.

Gehen wir von zwei Kausalzusammenhängen aus, so gibt es drei grundlegende Strukturmodelle, die aus diesen beiden gebildet werden können: Ein Gemeinsame-Ursache Modell, bei dem eine Ursache zwei Effekte hervorruft, ein Gemeinsamer-Effekt Modell, bei dem zwei Ursachen zu einem Effekt führen und ein Kettenstrukturmodell, bei dem eine erste Ursache ein intermittierendes Ereignis bedingt, welches wiederum einen Effekt hervorruft. Wie zuvor konzentrieren sich die Experimente auf das Gemeinsame-Ursache und das Gemeinsamer-Effekt Modell. Welche strukturellen Implikationen haben diese beiden Modelle für die Kovariation der beiden nicht unmittelbar kausal verbundenen Ereignisse? Die Gemeinsame-Ursache Struktur impliziert eine Kovariation der beiden Effekte, wobei die Höhe der Korrelation von der Stärke der beiden beteiligten Kausalzusammenhänge abhängig ist. Dabei gilt: Je häufiger die Ursache zu den beiden Effekten führt, desto stärker korrelieren diese miteinander. Die Gemeinsamer-Effekt Struktur impliziert keine Kovariation der beiden Ursachen. Es ist daher anzunehmen, dass Personen von Unabhängigkeit ausgehen, solange keine gegenteiligen Evidenzen vorliegen. Abbildung 3.32 gibt die Formeln zur Bestimmung der zu erwartenden Zusammenhänge an.

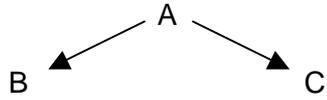
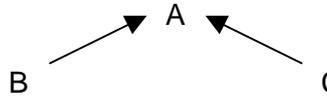
	Gemeinsame-Ursache Modell	Gemeinsamer-Effekt Modell
		
Wahrscheinlichkeitsverteilung	$P(B,C) = P(B A) \cdot P(C A) \cdot P(A) + P(B \sim A) \cdot P(C \sim A) \cdot P(\sim A)^{53}$	$P(B,C) = P(B) \cdot P(C)$
Zu erwartende Korrelation	$\Phi = [P(c.b) \cdot P(\sim c.\sim b) - P(c.\sim b) \cdot P(\sim c.b)] / [P(c) \cdot P(\sim c) \cdot P(b) \cdot P(\sim b)]$	$\Phi = [P(c) \cdot P(b) \cdot P(\sim c) \cdot P(\sim b) - P(c) \cdot P(b) \cdot P(\sim c) \cdot P(\sim b)] / [P(c) \cdot P(\sim c) \cdot P(b) \cdot P(\sim b)] = 0$

Abb. 3.32: Strukturmodelle und zu erwartende Korrelationen bei den Experimenten 6-8

⁵³ Diese Formel ergibt sich über den Satz der absoluten Wahrscheinlichkeit. Siehe auch Kapitel 1.4.

Als Maß für den Zusammenhang wurde dabei die Korrelation gewählt, da diese im Gegensatz zur Kontingenz stets unabhängig von der Richtung der Verbindung ist. Die dabei verwendeten Formeln sind äquivalent zur Bestimmung der 4-Felder bzw. Φ -Korrelation. Die Ausgangshypothese für die folgenden Experimente war, dass die Probanden sensitiv für die strukturellen Implikationen der Kausalmodelle sind. Sie sollten daher bei einem Gemeinsame-Ursache Modell einen Zusammenhang und bei einem Gemeinsamer-Effekt Modell eine Unabhängigkeit zwischen den beiden nicht direkt kausal verbundenen Ereignissen erwarten.

Die Sensitivität für strukturelle Implikationen wurde in den folgenden drei Experimenten jeweils mit zwei verschiedenen Maßen erhoben. Zur Messung der impliziten Sensitivität wurden die Probanden gebeten, auf der Grundlage des Modells Vorhersagen für neue Einzelfälle zu machen. So bekamen sie beim Gemeinsame-Ursache Modell neue Fälle gezeigt, in denen die Ursache entweder vorlag oder nicht gegeben war, und sie sollten angeben, ob die Effekte der Ursache ebenfalls vorlagen oder nicht. Beim Gemeinsamer-Effekt Modell erhielten sie Informationen über das Vorliegen des Effektes und sie sollten angeben, ob die Ursachen vorlagen oder nicht. Bei jeder Vorhersage nannten die Probanden also eine Kombination von Ereignissen. Über eine Folge von Einzelvorhersagen hinweg generierten die Probanden so Zusammenhänge zwischen den nicht kausal verbundenen Ereignissen. Dieses Maß setzt kein explizites Wissen über die strukturellen Implikationen von Kausalmodellen voraus. Für die Einzelvorhersagen wird lediglich ein prozedurales Wissen darüber verlangt, wie auf der Grundlage eines Kausalmodells Prognosen (Gemeinsame-Ursache Modell) und Diagnosen (Gemeinsamer-Effekt Modell) durchgeführt werden. Wie die Befunde in Kapitel 2.2 gezeigt haben, verfügen Probanden über diese Kompetenz. Zur Messung der expliziten Sensitivität für Kausalmodelle wurden die Probanden gebeten, für die einzelnen Modelle den Zusammenhang zwischen den nicht kausal verbundenen Ereignissen direkt einzuschätzen. Dieses Maß setzt ein explizites, deklaratives Wissen über die strukturellen Implikationen verschiedener Kausalmodelle voraus. Ob Probanden hierüber verfügen ist bisher noch unbekannt.

3.3.1 Experiment 6

In diesem Experiment wurde untersucht, ob Probanden sensitiv für die aus einem Strukturmodell folgenden Implikationen für den Zusammenhang nicht direkt kausal verbundener Ereignisse sind. Dazu wurden alle Probanden gebeten, zwei Kausalzusammenhänge unabhängig voneinander zu lernen. Die beiden Kausalzusammenhänge verbanden drei Ereignisse miteinander. Diese waren die Mutation eines Gens, das Auftreten eines Proteins und das Auftreten eines Enzyms. Alle Probanden lernten einerseits den Zusammenhang zwischen der Mutation und dem Auftreten des Proteins und andererseits den Zusammenhang zwischen der Mutation und dem Auftreten des Enzyms. Durch das separate Lernen wurde sichergestellt, dass die Probanden niemals eine Kovariation zwischen dem Auftreten des Proteins und des Enzyms beobachten konnten. Zwei experimentelle Faktoren wurden manipuliert. Als erster Faktor wurden die Annahmen über die den beiden Zusammenhängen zu Grunde liegende Kausalstruktur variiert. In einer Bedingung wurde den Probanden mitgeteilt, dass die beiden Kausalzusammenhänge ein Gemeinsame-Ursache Modell bilden, mit der Mutation als Ursache für das Auftreten des Protein und des Enzyms. In der anderen Bedingung wurden sie instruiert, dass den beiden Zusammenhängen eine Gemeinsamer-Effekt Struktur zu Grunde liegt, bei der die Mutation der Effekt der beiden Ursachen Enzym und Protein ist. Abbildung 3.33 stellt den Grundaufbau des Experimentes dar.

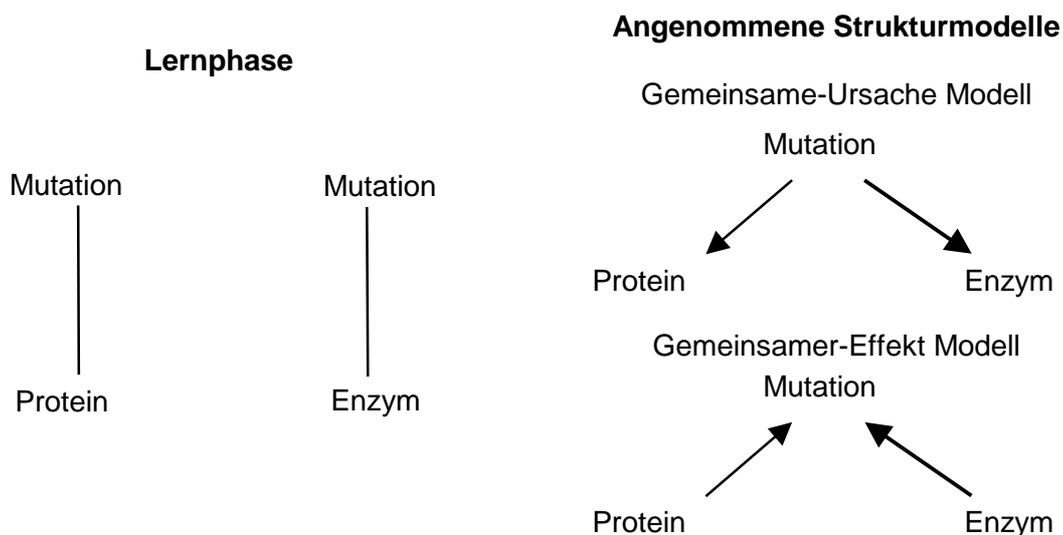


Abb. 3.33: Grundaufbau Experiment 6

Als zweiter Faktor wurde die Stärke der beiden beteiligten Kausalzusammenhänge variiert. Diese waren entweder beide stark, d.h., es lag eine hohe Kontingenz vor, oder beide waren schwach.

Um zu erfassen, ob Probanden sensitiv für die Implikationen der beiden Strukturen sind, wurden die oben eingeführten beiden Maße erhoben. Das explizite Wissen über den zu erwartenden Zusammenhang zwischen den beiden Ursachen bzw. den beiden Effekten wurde durch direkte Fragen erfasst. So wurden die Probanden gebeten, über die Einschätzung bedingter Wahrscheinlichkeiten die Stärke des nicht beobachteten Zusammenhangs zwischen den beiden Substanzen anzugeben. Um das implizite Wissen zu erfassen, wurden Probanden gebeten, für neue Fälle Vorhersagen zu machen. So wurden ihnen weitere Fälle gezeigt, in denen eine Mutation vorlag oder nicht vorlag. Die Probanden sollten angeben, ob in dem jeweiligen Fall das Protein und/oder das Enzym ebenfalls vorlag oder nicht vorlag. Über eine Reihe von Einzelvorhersagen generierten die Probanden so einen Zusammenhang zwischen den beiden Substanzen.

Folgende Vorhersagen ergeben sich für die vier Versuchsbedingungen. Wenn die Probanden sensitiv für die strukturelle Implikationen der verschiedenen Modelle sind, dann sollten sie bei einem Gemeinsamer-Effekt Modell keine Korrelation zwischen den beiden Substanzen vermuten, unabhängig davon, wie stark die beiden konstituierenden Kausalzusammenhänge sind. Bei einem Gemeinsame-Ursache Modell sollten sie bei den schwachen Zusammenhängen eine sehr schwache Korrelation annehmen, bei den starken Zusammenhängen aber eine hohe Korrelation.

Probanden und Design

48 Studenten der Universität Göttingen nahmen an dem Experiment teil. Sie erhielten für ihre Teilnahme wahlweise eine Versuchspersonenstunde testiert oder DM 7,- ausbezahlt. Die Probanden wurden per Zufall einer der vier Versuchsbedingungen zugewiesen. Das Experiment wurde im Einzelversuch durchgeführt.

Versuchsdurchführung und Material

Die Probanden in der Gemeinsame-Ursache Bedingung wurden instruiert, dass in letzter Zeit mehrfach beobachtet worden sei, dass das Gen 248 bei der Grünalge ‚arvor aquaticum‘ mutiert habe. Was diese Mutation alles bewirke, sei allerdings noch unklar. Wissenschaftler an verschiedenen biologischen Instituten in Deutschland seien an der Forschung beteiligt. So werde in Hamburg untersucht, ob die Mutation zu einer Bildung des Enzyms BST (BioSynergesTetrat) führe. Forscher am biologischen Institut der Universität Bremen hätten untersucht, ob die Mutation eine Bildung von Brasusproteinen hervorrufe. Die Aufgabe des Probanden werde sein hinterher einzuschätzen, wie sich die Mutation auf die

Bildung des Enzyms BST und des Brasusproteins auswirke. Dabei sei jedes Ergebnis denkbar. Es sei möglich, dass die Mutation zu einer häufigeren oder einer selteneren Bildung der beiden Substanzen führe oder aber die Bildung gar nicht beeinflusse. Was nun wirklich zutreffe, solle der Proband anhand der ihm gezeigten Karten herausfinden.

Die Probanden in der Gemeinsamer-Effekt Bedingung lasen in ihrer Instruktion, dass in letzter Zeit mehrfach beobachtet worden sei, dass das Gen 248 bei der Grünalge ‚arvor aquaticum‘ mutiert habe. Wovon diese Mutation bewirkt werde, sei allerdings noch unklar. Wissenschaftler an verschiedenen biologischen Instituten in Deutschland seien an der Forschung beteiligt. So werde in Hamburg untersucht, ob das Enzym BST (BioSynergesTetrat) zu einer Mutation führe. Forscher am biologischen Institut der Universität Bremen hätten untersucht, ob Brasusproteine eine Mutation hervorrufen. Die Aufgabe des Probanden werde sein, hinterher einzuschätzen, wie sich das Enzym BST und das Brasusprotein auf die Mutation des Gens 248 auswirken. Dabei sei jedes Ergebnis denkbar. Es sei möglich, dass die beiden Substanzen jeweils häufiger oder seltener zu einer Mutation führen oder aber die Mutation gar nicht beeinflussen. Was nun wirklich zutreffe, solle der Proband anhand der Karten herausfinden.

Die Probanden wurden gebeten, die Instruktion mit eigenen Worten zusammenzufassen und die vermutete Kausalstruktur aufzuzeichnen, um das Instruktionsverständnis zu überprüfen. Danach erhielten sie zwei Stapel mit je 40 Karten, welche die Daten der beiden Forschergruppen enthielten. Auf der Vorderseite der Karten erfuhren die Probanden, ob in dem spezifischen Fall eine Mutation vorlag, auf der Rückseite, ob das Enzym bzw. das Protein vorhanden waren oder nicht. Die Probanden blätterten selbständig die Karten durch, wobei sie nach jeder Karte zwischen den Stapeln zu wechseln hatten. Notizen machen oder Zurückblättern waren nicht erlaubt. Diese Prozedur war identisch für alle Probanden.

Nachdem sie die Daten durchgegangen waren, wurde das implizite Wissen über die strukturellen Implikationen der beiden Strukturen erfasst. Dazu wurden die Probanden gebeten, sich vorzustellen, dass sie Diplomand am biologischen Institut der Universität Göttingen seien und sich im Rahmen ihrer Diplomarbeit mit Mutationen bei der Grünalge ‚arvor aquaticum‘ beschäftigten. Sie hätten gerade 20 Algen daraufhin untersucht, ob ihr Gen 248 mutiert sei. Ihre Aufgabe sei es nun, für diese 20 Algen vorherzusagen, ob bei ihnen das Enzym BST und/oder das Brasusprotein vorliege. Das hieße, sie sollten sowohl für das Enzym BST als auch für das Brasusprotein eine Aussage machen. Den Probanden wurden daraufhin in zufälliger Reihenfolge 20 neue Fälle gezeigt, von denen die eine Hälfte eine Mutation aufwies und die andere nicht. Zu ihren Antworten erhielten die Probanden keine Rückmeldung.

Daraufhin wurde ihnen zwei Fragen vorgelegt, um ihr explizites Wissen erfassen. So wurden sie gefragt, wie häufig bei Vorliegen des Enzyms auch das Protein vorliegt und wie häufig bei Abwesenheit des Enzyms das Protein vorliegt. Die Fragerichtung (Enzym → Protein, bzw. Protein → Enzym) wurde dabei über die Teilnehmenden hinweg ausbalanciert. Ihre Antworten gaben die Probanden auf einer Skala von 0 (das Protein [Enzym] liegt nie vor) bis 100 (das Protein [Enzym] liegt immer vor).

Präsentierte Daten

Die gezeigten Zusammenhänge zwischen der Mutation und den Substanzen waren entweder beide stark oder beide schwach. Abbildung 3.34 gibt einen Überblick über die gezeigten Daten. Die angegebenen Häufigkeiten gelten sowohl für das Enzym wie für das Protein.

	Starke Zusammenhänge		Schwache Zusammenhänge	
	Enzym (Protein)	Kein Enzym (kein Protein)	Enzym (Protein)	Kein Enzym (kein Protein)
Mutation	16	4	10	10
Keine Mutation	0	20	6	14

Abb. 3.34: Daten zu Experiment 6

Wird ein Gemeinsame-Ursache Modell mit der Mutation als Ursache zu Grunde gelegt, so wird die Stärke der beiden Kausalzusammenhänge über $\Delta P = P(\text{Substanz}|\text{Mutation}) - P(\text{Substanz}|\text{keine Mutation})$ bestimmt. Daraus resultiert ein $\Delta P = 0,80$ für die starken Zusammenhänge, und $\Delta P = 0,20$ für die schwachen Zusammenhänge. Bei Annahme eines Gemeinsamer-Effekt Modells sind das Protein und das Enzym die Ursachen. Daher sind die Daten in entgegengesetzter Richtung zu analysieren, $\Delta P = P(\text{Mutation}|\text{Substanz}) - P(\text{Mutation}|\text{keine Substanz})$. Es ergeben sich sehr ähnliche Kontingenzen wie zuvor, $\Delta P = 0,83$ für die starken und $\Delta P = 0,21$ für die schwachen Zusammenhänge.

Welche strukturellen Implikationen ergeben sich aus den Modellen bei diesen Daten für den Zusammenhang zwischen dem Protein und dem Enzym? Mittels der oben vorgestellten Gleichungen können die zu erwartenden Korrelationen berechnet werden. Ein Gemeinsamer-Effekt Modell impliziert keine Kovariation. Deshalb sollte der Zusammenhang unabhängig von der Stärke der beiden Kausalzusammenhänge Null sein. Die von dem Gemeinsame-Ursache Modell und der Stärke der Kausalbeziehungen implizierten Korrelationen zwischen den Substanzen sind $\Phi = 0,67$ für die starken Kausalrelationen und

$\Phi = 0,04$ für die schwachen Kausalrelationen. Abbildung 3.35 gibt nochmals einen Überblick über alle gezeigten und von den jeweiligen Modellen implizierten Zusammenhänge.

	Starke Zusammenhänge	Schwache Zusammenhänge
	<p>Mutation $\Delta P = 0,80$ $\Delta P = 0,83$ Enzym Protein</p>	<p>Mutation $\Delta P = 0,20$ $\Delta P = 0,21$ Enzym Protein</p>
Implizierte Korrelation Gemeinsame-Ursache Modell	Φ (Enzym,Protein) = 0,67	Φ (Enzym,Protein) = 0,04
Implizierte Korrelation Gemeinsamer-Effekt Modell	Φ (Enzym,Protein) = 0,00	Φ (Enzym,Protein) = 0,00

Abb. 3.35: Beobachtbare und implizierte Zusammenhänge in Experiment 6

Ergebnisse

Als implizites Maß wurde die über die Einzelvorhersagen generierte Φ -Korrelation der beiden Substanzen verwendet. Um auch für das explizite Maß denselben Maßstab zu haben, wurden die Antworten auf die expliziten Fragen ebenfalls in Φ -Korrelationen umgerechnet. Hierfür wurden die geschätzten bedingten Wahrscheinlichkeiten in die Gleichungen aus Abbildung 3.32 eingesetzt. Die Mittelwerte für die generierte wie für die geschätzte Korrelation der beiden Substanzen zeigt Abbildung 3.36.

	Implizites Maß: Generierte Korrelation		Explizites Maß: Geschätzte Korrelation	
	Gemeinsame- Ursache Modell	Gemeinsamer- Effekt Modell	Gemeinsame- Ursache Modell	Gemeinsamer- Effekt Modell
Starke Zusammenhänge	0,622	0,168	0,286	0,161
Schwache Zusammenhänge	-0,004	-0,130	-0,109	0,039

Abb. 3.36: Ergebnisse Experiment 6: Mittelwerte der generierten und der geschätzten Korrelation zwischen Enzym und Protein

Die von den Probanden generierten Korrelationen kommen den implizierten Korrelationen erstaunlich nahe. Bei Annahme eines Gemeinsame-Ursache Modells wurde bei schwachen Zusammenhängen eine geringe ($\Phi = -0,004$) und bei starken Zusammenhängen eine hohe Korrelation ($\Phi = 0,62$) zwischen den Substanzen generiert.

Nahmen die Probanden aber ein Gemeinsamer-Effekt Modell an, so erzeugten sie trotz identischer Daten in beiden Bedingungen eine geringe Korrelation. Eine Varianzanalyse für die generierten Korrelationen mit Stärke des Zusammenhangs und Strukturmodell als Zwischensubjektfaktoren erbrachte sowohl einen signifikanten Haupteffekt für Strukturmodell, $F_{Struktur}(1,44)= 7,28, p < 0,05, MSE = 0,14$, als auch für Zusammenhangsstärke, $F_{Stärke_Zshg.}(1,44)= 18,4, p < 0,01, MSE = 0,14$. Die Interaktion wurde nicht signifikant, $F(1,44)= 2,33, p = 0,13, MSE = 0,14$.

Die geschätzten Korrelationen wichen erheblich von den implizierten Zusammenhängen ab. Eine Varianzanalyse ergab lediglich einen signifikanten Effekt für die Stärke des Kausalzusammenhangs, $F_{Stärke_Zshg.}(1,44)= 8,05, p < 0,01, MSE = 0,10$. Es zeigte sich kein Unterschied zwischen den beiden Strukturen ($F < 1$) und auch die Interaktion wurde nicht signifikant.⁵⁴

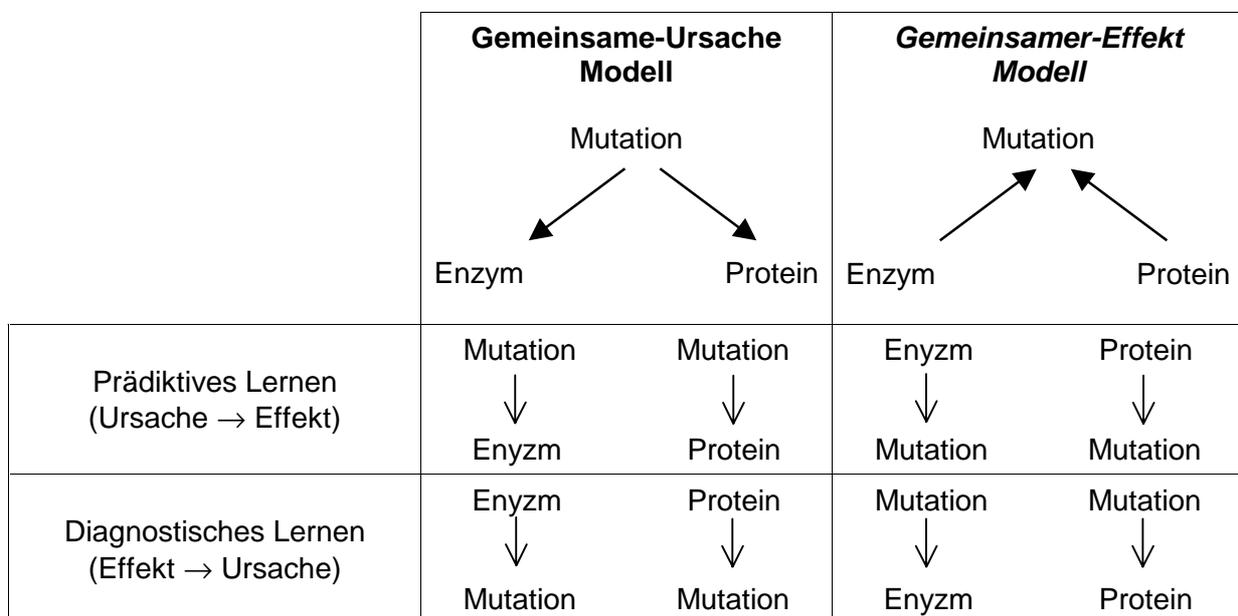
Die Ergebnisse deuten darauf hin, dass die Probanden generell bei stärkeren Kausalzusammenhängen höhere Korrelationen erwarteten. Dieser Befund zeigte sich für die generierten wie für die geschätzten Korrelationen. Die Sensitivität für strukturelle Implikationen war aber auf das implizite Maß beschränkt. Nur die mit den Einzelaussagen generierten Korrelationen stimmten mit den normativ zu erwartenden überein. Dagegen wurden die expliziten Einschätzungen von der angenommenen Kausalstruktur nicht beeinflusst.

3.3.2 Experiment 7

In Experiment 6 wurden die Probanden stets zuerst darüber informiert, ob eine Mutation stattgefunden hatte und dann unabhängig voneinander, ob ein Enzym oder ein Protein gefunden wurde. Damit war der Lerninput für die Probanden in allen Bedingungen derselbe. Allerdings bedingt dieses identische Vorgehen auch eine Asymmetrie. Bei Annahme eines Gemeinsame-Ursache Modells bedeutet diese Prozedur, dass die Probanden zunächst über die Ursache (Mutation) und dann über den Effekt (Substanz) informiert wurden. Sie lernten also prädiktiv. Dagegen wird unter Zugrundelegung einer Gemeinsamer-Effekt Struktur vom Effekt (Mutation) auf die Ursache (Substanz) gelernt. Damit lernten die Probanden in dieser Bedingung diagnostisch. Frühere Untersuchungen haben gezeigt, dass es Unterschiede zwischen prädiktivem und diagnostischem Lernen gibt

⁵⁴ Da die generierten und die geschätzten Korrelationen auf unterschiedlichen Maßen beruhen, wurde keine gemeinsame Analyse durchgeführt. So gehen in die generierte Korrelation 20 Einzelantworten ein, in die geschätzte Korrelation dagegen nur zwei Antworten. Die Reliabilität der generierten Korrelation dürfte daher höher sein.

(Reips, 1997; Waldmann, 1996).⁵⁵ Ausgehend von dieser unterschiedlichen Bedeutung des Inputs stellt sich die Frage, ob der Befund aus Experiment 6 auf die Unterschiede im Lernen zurückgeführt werden kann. So könnte man spekulieren, dass bei einem Gemeinsamer-Effekt Modell deshalb kein Zusammenhang zwischen den Ursachen generiert wird, weil das Lernen wegen der höheren Schwierigkeit noch nicht so weit fortgeschritten ist wie beim prädiktiven Lernen. Ziel von Experiment 7 ist es daher, den Befund des vorherigen Experimentes zu replizieren und dabei den möglichen Einfluss der Lernrichtung zu kontrollieren. Die Aufgabe der Probanden war es erneut, zwei Kausalzusammenhänge getrennt voneinander zu erlernen. Als Hintergrundgeschichte wurde wie bei dem Vorgängerexperiment der Zusammenhang zwischen einer Genmutation und dem Vorliegen eines Proteins und eines Enzyms verwendet. Manipuliert wurden zwei Faktoren. Zunächst wurde wiederum die Annahme über die zu Grunde liegende Kausalstruktur variiert. So gingen die Versuchsteilnehmer entweder von einem Gemeinsame-Ursache Modell aus, mit der Mutation als Ursache der beiden Substanzen, oder von einem Gemeinsamer-Effekt Modell, mit der Mutation als Effekt der Substanzen. Als zweiter Faktor wurde die Lernrichtung variiert. Die Probanden lernten entweder prädiktiv, von der vermeintlichen Ursache auf den Effekt, oder diagnostisch, vom Effekt auf die Ursache. Nicht verändert wurde die Stärke der beteiligten Kausalzusammenhänge. Diese waren konstant von mittlerer Stärke. Alle Probanden erhielten exakt dieselben Daten. Abbildung 3.37 stellt die vier Bedingungen im Überblick dar.



*Abb. 3.37: Versuchsbedingungen Experiment 7.
Die Pfeile in den einzelnen Zellen geben die Lernrichtung an.*

⁵⁵ So ist beispielsweise prädiktives Lernen häufig einfacher, während dagegen diagnostisches Lernen sensitiver für Unterschiede in den Basisraten verschiedener Ursachen ist.

Welche Ergebnisse waren zu erwarten? Wie in Experiment 6 sollten die Probanden bei Annahme eines Gemeinsame-Ursache Modells eine Korrelation zwischen den Substanzen vermuten, bei einem Gemeinsamer-Effekt Modell aber die Substanzen für unabhängig halten. Es ist zu erwarten, dass sich dieser Effekt wieder nur im impliziten, nicht aber im expliziten Maß zeigt. Ob die Lernrichtung einen Einfluss auf den Effekt hat, müssen die empirischen Befunde zeigen.

Probanden

64 Studierende der Universität Göttingen nahmen an dem Experiment teil. Sie erhielten für ihre Teilnahme wahlweise DM 10,-- oder eine Versuchspersonenstunde testiert. Die Zuteilung zu einer der vier Versuchsbedingungen erfolgte per Zufall. Das Experiment wurde im Einzelversuch durchgeführt.

Versuchsdurchführung und Material

Die Prozedur wurde von Experiment 6 übernommen. Alle Probanden hatten die Aufgabe zwei Kausalzusammenhänge getrennt voneinander zu erlernen. In der Gemeinsame-Ursache Bedingung wurden sie erneut instruiert, dass von einer Forschergruppe untersucht werde, ob die Mutation des Gens 248 die Bildung des Enzyms BST hervorrufe. Eine andere Forschergruppe untersuche, ob die Mutation zur Bildung von Brasusproteinen führe. In der Gemeinsamer-Effekt Bedingung wurde wiederum instruiert, dass eine Gruppe von Forschern die Auswirkung des Enzyms BST auf die Mutation des Gens und die Forscher einer anderen Gruppe die Wirkung von Brasusproteinen auf die Mutation untersuchen würden. Nach der Instruktion und der Überprüfung des Instruktionsverständnisses erhielten die Probanden zwei Stapel mit je 50 Karteikarten, welche die Untersuchungsbefunde der beiden Forschergruppen beinhalteten. Zur Manipulation der Lernrichtung wurden die Kartenstapel entsprechend umgedreht. So lag beispielsweise für prädiktives Lernen in der Gemeinsame-Ursache Bedingung die Seite mit der Mutation oben. Durch Umdrehen erfuhr der Proband, ob das Enzym bzw. das Protein vorlag oder nicht. Für diagnostisches Lernen beim Gemeinsame-Ursache Modell wurden die Karten anders herum gezeigt, d.h. die Seite mit der Substanz lag oben. Beim Gemeinsamer-Effekt Modell wurde entsprechend verfahren.

Nachdem die Probanden abwechselnd die beiden Kartenstapel studiert hatten, wurde ihr implizites Wissen erfasst. Sie wurden gebeten für 20 neue Fälle Vorhersagen zu machen. In der Hälfte dieser Fälle lag eine Mutation vor, in der anderen nicht. Die Probanden sollten angeben, ob in dem jeweiligen Fall das Enzym und/oder das Protein vorlag oder nicht. Sie erhielten auf ihre Antworten keinerlei Feedback. Danach wurden sie explizit nach dem erwarteten Zusammenhang der beiden Substanzen befragt. Wie in Experiment 6 sollten sie

angeben, wie häufig das Protein bei Anwesenheit des Enzyms und wie häufig es bei Abwesenheit des Enzyms vorlag. Die Fragerichtung wurde erneut ausbalanciert. Die Antwortskala von 0 (nie) bis 100 (immer) wurde ebenfalls übernommen.

Präsentierte Daten

Die Daten waren für alle Probanden identisch. Zudem waren die Daten symmetrisch gestaltet, so dass die bedingten Wahrscheinlichkeiten sowohl von den Substanzen auf die Mutation, als auch von der Mutation auf die Substanzen identisch waren.⁵⁶ Abbildung 3.38 zeigt die verwendeten Häufigkeiten.

	Enzym (Protein)	Kein Enzym (kein Protein)
Mutation	20	5
Keine Mutation	5	20

Abb. 3.38: Daten zu Experiment 7

Daraus ergeben sich folgende beobachtbare und implizierte Zusammenhänge.

Kausalzusammenhänge	<pre> graph TD Mutation -- "ΔP = 0,60" --> Enzym Mutation -- "ΔP = 0,60" --> Protein Enzym -.-> Protein </pre>
Implizierte Korrelation Gemeinsame-Ursache Struktur	$\Phi(\text{Enzym, Protein}) = 0,36$
Implizierte Korrelation Gemeinsamer-Effekt Struktur	$\Phi(\text{Enzym, Protein}) = 0,00$

Abb. 3.39: Beobachtbare und implizierte Zusammenhänge in Experiment 7

Ergebnisse

Mit ihren Einzelvorhersagen generierten die Probanden eine Korrelation zwischen den Substanzen. Diese wurde als implizites Maß für ihre Sensitivität verwendet. Die von den Probanden geschätzten bedingten Wahrscheinlichkeiten zwischen den Substanzen wurden wieder in Φ -Korrelationen umgerechnet. Diese geben die explizite Schlussfolgerung der Probanden über den Zusammenhang der beiden Substanzen wieder. Abbildung 3.40 zeigt die Mittelwerte sowohl für das implizite wie das explizite Maß.

⁵⁶ Damit waren sowohl die Kontingenzen als auch die Werte für die kausale Power in beiden Richtungen identisch.

	Implizites Maß: Generierte Korrelation		Explizites Maß: Geschätzte Korrelation	
	Gemeinsame- Ursache Modell	Gemeinsamer- Effekt Modell	Gemeinsame- Ursache Modell	Gemeinsamer- Effekt Modell
Prädiktives Lernen	0,243	0,013	0,258	0,239
Diagnostisches Lernen	0,186	-0,001	0,129	0,210

Abb. 3.40: Ergebnisse Experiment 7: Mittelwerte der generierten und der geschätzten Korrelation zwischen Enzym und Protein

Die generierten Korrelationen zeigen einen deutlichen Unterschied zwischen den beiden Strukturmodellen. Eine Varianzanalyse mit Lernrichtung und Strukturmodell als Zwischensubjektfaktoren ergab nur einen signifikanten Haupteffekt für das angenommene Kausalmodell, $F(1,62) = 4,97$, $p < 0,05$, $MSE = 0,14$. Bei Annahme eines Gemeinsame-Ursache Modells erzeugten die Probanden eine Korrelation, die deutlich größer als Null war. Dagegen hielten sie bei einem Gemeinsamer-Effekt Modell die beiden Substanzen für unabhängig. Die Lernrichtung hatte keinen statistisch signifikanten Einfluss auf die generierte Korrelation und interagierte auch nicht mit dem gezeigten Kausalmodell ($F < 1$ jeweils).

Für die explizit geschätzte Korrelation ergab eine Varianzanalyse keinerlei signifikanten Effekte (alle $F < 1$). Die Probanden gingen stets davon aus, dass ein gewisser Zusammenhang der beiden Substanzen vorlag. Das angenommene Kausalmodell hatte hierauf keinen Einfluss, ebenso wenig wie die Lernrichtung.

Damit repliziert Experiment 7 die Ergebnisse des vorherigen Experimentes. Zusätzlich belegt es, dass die implizite strukturelle Sensitivität unabhängig davon ist, ob die Kausalzusammenhänge diagnostisch oder prädiktiv gelernt werden.

3.3.3 Experiment 8

In den beiden vorherigen Experimenten wurde jeweils sichergestellt, dass die Probanden keinen Zusammenhang zwischen den beiden Effekten bzw. Ursachen beobachten konnten. Ihre Einzelvorhersagen wie ihre Einschätzungen mussten sie deshalb vollständig auf ihr implizites bzw. explizites Verständnis der Implikationen verschiedener Strukturmodelle stützen. Nun stellt sich die Frage, was passiert, wenn der Zusammenhang der beiden vermeintlichen Effekte bzw. Ursachen beobachtet werden kann. Experiment 8 untersucht daher zwei Faktoren. Als erstes wurde erneut die hinter den beiden gelernten Kausalzusammenhängen vermutete Struktur variiert. Wie zuvor gingen die Probanden entweder von einem Gemeinsame-Ursache Modell oder einem Gemeinsamer-Effekt Modell aus. Als zweiter Faktor wurde manipuliert, ob die Probanden eine Korrelation zwischen den Ursachen bzw. Effekten beobachten konnten oder nicht. In der Nichtbeobachtungsbedingung wurden ihnen wie bisher die Daten verteilt auf zwei unabhängigen Kartenstapeln gezeigt. In der Beobachtungsbedingung wurde ihnen lediglich ein Stapel gezeigt, wobei die Probanden gleichzeitig Informationen über das Vorliegen der Mutation, des Enzyms und des Proteins erhielten. Die beobachtbare Korrelation zwischen den beiden Substanzen entsprach dabei derjenigen, die von dem Gemeinsame-Ursache Modell impliziert wird. Die Aufgabe in allen Bedingungen war es, die Stärke der Kausalzusammenhänge zu ermitteln.

Welche Ergebnisse sind zu erwarten? Die Nichtbeobachtungsbedingung ist eine weitere Replikation von Experiment 6. Die Probanden sollten daher bei ihren Einzelvorhersagen erneut eine Sensitivität für die strukturellen Implikationen der beiden Strukturmodelle zeigen.

Die Beobachtungsbedingung bietet nun im Gegensatz zur Nichtbeobachtungsbedingung direkt Daten bezüglich der implizierten Korrelation an. Den Probanden liegen also zwei Quellen für ihre Einschätzung vor, einerseits die gesehene Korrelation, andererseits die auf Grundlage des Strukturmodells und der Stärke der beobachteten Kausalzusammenhänge zu erwartende Korrelation. Sofern die Probanden also neben den beiden Kausalzusammenhängen auch Wissen über die Kovariation der beiden Substanzen erwerben, sollten sie in der Lage sein, diese richtig zu reproduzieren und auch Fragen in Bezug auf diese richtig zu beantworten. Bisherige Forschungsarbeiten haben gezeigt, dass Probanden bei einfachen Aufgaben Wissen über die Zusammenhänge zwischen allen beteiligten Ereignissen erwerben können, unabhängig davon, ob es sich um kausale, implizierte oder zufällige Zusammenhänge handelt (vgl. Waldmann, 1994, für einen Überblick). Es ist daher zu erwarten, dass die Probanden in der Beobachtungsbedingung in der Lage sind, auch bei den expliziten Fragen zu einer systematisch richtigen Antwort zu kommen. Diese Antworten sollten sich auf die beobachtete Korrelation beziehen.

Der beobachtete Zusammenhang der beiden Substanzen widerspricht jedoch den Implikationen des Gemeinsamer-Effekt Modells. Dieses Modell impliziert ja eine Unabhängigkeit der beiden Ursachen. Was ist in diesem Fall zu erwarten? Dazu ist zu sagen, dass das Gemeinsamer-Effekt Modell zwar eine Nullkorrelation impliziert, aber eine höhere Korrelation nicht ausschließt. Es lässt zu, dass die Ursachen kontingenterweise kovariieren. Diese Kovariation ist dabei keine Folge der Ereignisse innerhalb des Modells, sondern unbekannter Faktoren, welche im Modell fehlen. Daher ist zu erwarten, dass die Probanden ihre Einschätzungen auf die beobachtete Korrelation und nicht auf die implizierte Unabhängigkeit stützen.

Zusammenfassend gesagt: Für die Nichtbeobachtungsbedingung werden die gleichen Befunde wie in den Experimenten zuvor erwartet. Bei den Einzelvorhersagen sollte sich ein Sensitivität für strukturelle Implikationen zeigen, bei den expliziten Einschätzungen nicht. In der Beobachtungsbedingung sind für das explizite und das implizite Maß gleich hohe Korrelationen zu erwarten. Beide sollten der beobachtbaren Korrelation zwischen den Substanzen entsprechen.

Probanden und Design

Es nahmen 56 Studierende der Universität Göttingen teil. Sie erhielten entweder eine Versuchspersonenstunde testiert oder DM 10,--. Keine der Versuchspersonen hatte vorher an einem anderen Experiment der Versuchsreihe teilgenommen. Die Probanden wurden per Zufall einer der vier Versuchsbedingungen zugeteilt. Der Versuch wurde einzeln durchgeführt.

Versuchsdurchführung und Material

Das Material wurde aus den beiden Vorgängerexperimenten übernommen. Wiederum handelte es sich um die Aufgabe, die Kausalzusammenhänge zwischen der Mutation eines Gens und einem Enzym und einem Protein zu lernen. Die Durchführung war ebenfalls dieselbe. Einziger Unterschied war die Darbietung der Daten in der Beobachtungsbedingung. In der Nichtbeobachtungsbedingung erhielten die Probanden wie bisher zwei Stapel mit je 40 Karten, welche jeweils die Untersuchungsbefunde einer Forschergruppe zeigten und vollständig unabhängig waren. Auf der Vorderseite erfuhren sie, ob eine Mutation vorlag, und auf der Rückseite, ob das Protein bzw. beim anderen Stapel das Enzym vorlag. In der Beobachtungsbedingung erhielten die Probanden einen Stapel mit 80 Karten. Auf der Vorderseite war zu lesen, ob eine Mutation vorlag und auf der Rückseite, ob die beiden Substanzen vorlagen. Nach dem Studium der Daten wurden wie zuvor das implizite und das explizite Maß für die erwartete Korrelation der beiden Substanzen erhoben.

Präsentierte Daten

Die gezeigten Kausalzusammenhänge waren beide von mittlerer Stärke. Wie in Experiment 7 wurde darauf geachtet, dass die bedingten Wahrscheinlichkeiten zwischen der Mutation und den Substanzen in beiden Richtungen identisch waren. Abb. 3.41 zeigt die verwendeten Häufigkeiten.

	Enzym (Protein)	Kein Enzym (kein Protein)
Mutation	17	3
Keine Mutation	3	17

Abb. 3.41: Daten zu Experiment 8

Daraus ergeben sich folgende beobachtbare und implizierte Zusammenhänge.

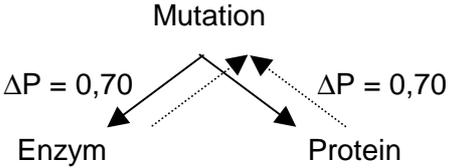
	Kausalzusammenhänge	
Korrelation nicht beobachtbar	Implizierte Korrelation Gemeinsame-Ursache Struktur	$\Phi(\text{Enzym, Protein}) = 0,49$
	Implizierte Korrelation Gemeinsamer-Effekt Struktur	$\Phi(\text{Enzym, Protein}) = 0,00$
Korrelation beobachtbar	Beide Strukturen	$\Phi(\text{Enzym, Protein}) = 0,50$

Abb. 3.42: Beobachtbare und implizierte Zusammenhänge in Experiment 8

Ergebnisse

Abbildung 3.43 fasst die Ergebnisse zusammen. Für die generierten Korrelationen zeigte sich erneut eine Sensitivität für strukturelle Implikationen. In der Nichtbeobachtungsbedingung wurde für das Gemeinsame-Ursache Strukturmodell eine hohe, für das Gemeinsamer-Effekt Modell eine Nullkorrelation erzeugt. War allerdings die Kovariation beobachtbar, so verschwand dieser Unterschied. In beiden Fällen wurde ein Zusammenhang erzeugt. Eine Varianzanalyse mit Strukturmodell und Beobachtbarkeit als Zwischensubjektfaktoren erbrachte zwei signifikante Haupteffekte, aber keine signifikante

Interaktion, $F_{Struktur}(1,52) = 5,02$, $p < 0,05$, $MSE = 0,13$; $F_{Beobachtbarkeit}(1,52) = 9,41$, $p < 0,01$, $MSE=0,13$; $F_{Interaktion}(1,52) = 1,46$, $p = 0,23$, $MSE = 0,13$.

	Implizites Maß: Generierte Korrelation		Explizites Maß: Geschätzte Korrelation	
	Gemeinsame- Ursache Modell	Gemeinsamer- Effekt Modell	Gemeinsame- Ursache Modell	Gemeinsamer- Effekt Modell
Korrelation nicht beobachtbar	0,300	-0,030	0,219	0,075
Korrelation beobachtbar	0,478	0,380	0,294	0,355

Abb. 3.43: Ergebnisse Experiment 8: Mittelwerte der generierten und der geschätzten Korrelation zwischen Enzym und Protein

Die expliziten Einschätzungen sehen oberflächlich betrachtet den impliziten ähnlich. Eine Varianzanalyse erbrachte aber keine signifikanten Effekte. Tendenziell zeigte sich lediglich ein Effekt für die Beobachtbarkeit, $F(1,52) = 3,34$, $p < 0,10$, $MSE = 0,13$. Zwischen den beiden Modellen wurde nicht unterschieden, ($F < 1$). Auch die Interaktion war nicht signifikant, $F(1,52) = 1,12$, $p = 0,30$, $MSE = 0,13$. Ein Einzelvergleich der Einschätzungen in der Nichtbeobachtungsbedingung ergab keinen signifikanten Effekt, obwohl zwischen den Mittelwerten scheinbar ein Unterschied besteht, $F(1,26) = 0,87$, $p = 0,36$, $MSE = 0,17$.

Die Ergebnisse bestätigen teilweise die gemachten Vorhersagen. Hatten die Probanden keine Informationen über den Zusammenhang der beiden Substanzen, so generierten sie die von dem jeweiligen Strukturmodell implizierten Korrelationen. Dieser Befund repliziert die Ergebnisse der beiden Vorgängerexperimente. Jedoch war diese Korrelation schwächer als die in der Beobachtungsbedingung generierte. Dies ist möglicherweise darauf zurückzuführen, dass die Probanden sich mit ihrem Vorhersagen sicherer waren, wenn sie zuvor den Zusammenhang beobachten konnten.

Bei den expliziten Urteilen zeigte sich erneut keine Sensitivität für strukturelle Implikationen.⁵⁷ Zusätzlich belegen die Einschätzungen in der Beobachtungsbedingung, dass die Probanden sich mit ihren Urteilen an der beobachtbaren Korrelation orientiert haben.

⁵⁷ Der rein deskriptive Unterschied zwischen den Mittelwerten in der Nichtbeobachtungsbedingung deutet aber darauf hin, dass unter Umständen bei höherer Power an Versuchspersonen sich auch hier ein Unterschied ergeben könnte.

Der letzte Befund mag nicht überraschend klingen, aber er widerspricht Theorien kausalen Denkens, die davon ausgehen, dass nur Kausalzusammenhänge gespeichert werden. In diesem Fall sollten die Probanden bei Annahme eines Gemeinsamer-Effekt Modells den Zusammenhang der beiden Ursachen unterschätzen. Dafür haben wir keine Belege gefunden. Allerdings ist der dargestellte Sachverhalt mit seinen drei Ereignissen sehr einfach. Drei Zusammenhänge zu lernen, ist für Personen leicht möglich, insbesondere, wenn der Zusammenhang zwischen den nicht kausal verbundenen Ereignissen, wie in dem vorliegenden Experiment, recht hoch ist. Es ist deshalb zu erwarten, dass die Probanden bei komplexeren Sachverhalten mit mehr Ereignissen sich tatsächlich nur die Kausalzusammenhänge einprägen. Sie sollten dann weitere, zufällige Zusammenhänge in den Daten nicht erkennen oder zumindest in Richtung des von der Struktur implizierten Zusammenhangs verzerren. Ob dies der Fall ist, muss zukünftige Forschung zeigen.

Diskussion

Die Fragestellung der Experimente 6 bis 8 war, ob Probanden bei der Integration von partiellem Wissen in Kausalmodelle sensitiv für deren strukturelle Implikationen sind. Um dies zu untersuchen, lernten die Probanden unabhängig voneinander zwei Kausalzusammenhänge. Es wurden die Annahmen über das den beiden Zusammenhängen zu Grunde liegende Kausalmodell manipuliert. Nach dem Lernen wurden die Probanden gebeten, Aussagen über den Zusammenhang der beiden Ursachen bzw. der beiden Effekte zu machen, ein Zusammenhang, den sie nie zuvor gesehen hatten. Um diese Aufgabe richtig zu lösen, war es notwendig, die strukturellen Implikationen der verschiedenen Modelle zu berücksichtigen. In allen drei Experimenten konnte eine Sensitivität mittels eines impliziten Maßes nachgewiesen werden. Die Probanden generierten über eine Folge von Vorhersagen für einzelne Fälle hinweg die Korrelation, die aufgrund der Modellstruktur und der Stärke der beteiligten Kausalzusammenhänge zu erwarten gewesen wäre. Für das Gemeinsamer-Effekt Modell wurde in allen drei Experimenten tendenziell eine Nullkorrelation erzeugt. Für das Gemeinsame-Ursache Modell zeigt sich über die drei Experimente hinweg ein interessanter Trend. Der generierte Zusammenhang steigt mit der Stärke der beteiligten Kausalzusammenhänge an. Abbildung 3.44 veranschaulicht diesen Trend.

	Stärke der Kausalzusammenhänge	Implizierte Korrelation	Generierte Korrelation
Experiment 1	$\Delta P = 0,2$	$\Phi = 0,04$	-0,004
Experiment 2	$\Delta P = 0,6$	$\Phi = 0,36$	0,243
Experiment 3	$\Delta P = 0,7$	$\Phi = 0,49$	0,300
Experiment 1	$\Delta P = 0,8$	$\Phi = 0,64$	0,622

Abb. 3.44: Generierte Korrelation der Effekte bei einem Gemeinsame-Ursache Modell in Abhängigkeit der Stärke der beteiligten Kausalzusammenhänge

Dies zeigt, dass die Probanden nicht nur zwischen schwachen und starken Kausalzusammenhängen einen Unterschied machen, wie Experiment 1 gezeigt hat, sondern die jeweilige Stärke mit in Betracht ziehen.

Es gibt Studien anderer Forscher (Ahn & Dennis, 2000), die eine Sensitivität für strukturelle Implikationen auch bei kausalen Ketten nachweisen konnten. Diese gaben Probanden die Aufgabe, zwei unabhängig voneinander gelernte Kausalzusammenhänge zu einer kausalen Kette zu verbinden. Die Probanden sollten daraufhin einschätzen, wie sehr die erste Ursache den letzten Effekt beeinflusst. Die Stärke des vermuteten Einflusses stieg mit der Stärke der beteiligten Kausalzusammenhänge an. Die Urteile stimmten mit den Vorhersagen überein, die sich unter der Annahme einer Markovkette (vgl. Kapitel 1.2) ergeben.

Ein zweiter Hauptbefund der obigen Experimente war, dass die Probanden bei ihren expliziten Urteilen keine Sensitivität für strukturelle Implikationen zeigen. Dies waren aber genau dieselben Probanden, welche zuvor die Korrelationen richtig generiert hatten. Wie kommt es zu diesem Unterschied zwischen dem impliziten und dem expliziten Maß? Wahrscheinlich haben die Probanden unterschiedliche Strategien für die Bearbeitung der jeweiligen Aufgabe eingesetzt. Bei der impliziten Aufgabe wurden den Probanden neue, bisher nicht gesehene Fälle vorgelegt. Dabei war das Gen bei einem Teil der Algen mutiert, bei anderen Algen nicht. Die Probanden sollten für jeden Einzelfall vorhersagen, ob ein Enzym und/oder ein Protein vorlagen. Eine solche Aufgabe kann sehr gut über mentale Simulation gelöst werden. Ist die Mutation die gemeinsame Ursache, dann wird das Auftreten der Ursache simuliert. Dadurch werden je nach Stärke der Kausalverbindungen die beiden Effekte mit einer gewissen Wahrscheinlichkeit hervorgerufen. Dies führt über mehrere Fälle hinweg einerseits zu dem bekannten Phänomen des ‚Probability matching‘ (u.a.

Davison & McCarthy, 1988) andererseits entsteht dabei (sozusagen als Nebeneffekt) die implizierte Korrelation der Effekte. Sie braucht also nicht eigens generiert werden. Ist die Mutation der gemeinsame Effekt, so werden mögliche Ursachenkombinationen aus dem Gedächtnis abgerufen. Zur Kontrolle, ob diese Kombination die beobachtete Mutation hervorgerufen hat, kann ebenfalls eine Simulation durchgeführt werden. Dies ist aber nicht notwendig (vgl. Simulationsheuristik, Kahneman & Tversky, 1982). Die Unabhängigkeit entsteht dadurch, dass mehrere verschiedene Ursachenkombinationen mit dem aufgetretenen Effekt vereinbar sind. Liegt die Mutation vor, so können sowohl beide als auch nur jeweils eine Substanz vorgelegen haben.

Dafür, dass die Probanden mentale Simulationen eingesetzt haben, gibt es einen direkten Hinweis. In ihren Kommentaren gaben viele Probanden an, dass sie sich bei der Vorhersageaufgabe den jeweiligen Fall einfach vorgestellt hätten. Weitere indirekte Hinweise ergeben sich aus anderen Forschungsarbeiten aus dem Bereich des kausalen Denkens. Mentale Simulation wird von Probanden bei verschiedenen Aufgaben eingesetzt, u.a. bei der Beantwortung kontrafaktischer Fragen (Roese, 1997; Spellman & Mandel, 1999; Wells & Gavanski, 1989). Ein Beispiel für eine solche Frage ist: Wäre die Concorde nicht abgestürzt, wenn der Reifen nicht geplatzt wäre? Um diese Frage zu beantworten, müssen sich die Personen einen Fall vorstellen, den sie nicht gesehen haben und der dem konkret beobachteten Fall entgegengesetzt ist. Unsere Aufgabe, Vorhersagen für neue Fälle zu machen, gleicht dieser Frage darin, dass ein noch nicht beobachteter Fall zu beurteilen ist. Sie weicht aber insofern davon ab, dass die Probanden bereits mehrere gleichartige Fälle zuvor gesehen hatten.

Wie kommen Probanden zu ihren expliziten Urteilen? Ist der zu beurteilende Zusammenhang nicht wie in Experiment 8 direkt beobachtbar, müssten die Probanden, um diese Aufgabe richtig zu lösen, Bayesianische Berechnungen anstellen. Dazu sind sie offensichtlich nicht in der Lage. Die expliziten Einschätzungen zeigten keine erkennbare Systematik über die Experimente hinweg, allenfalls, dass bei starken Kausalzusammenhängen unabhängig vom Kausalmodell ein höherer Zusammenhang vermutet wird (vgl. Experiment 6). Es scheint so, als ob die Probanden die Antwort auf die Fragen nach den bedingten Wahrscheinlichkeiten raten würden. In dieser Weise äußerten sich auch viele Probanden in ihren Kommentaren zu diesen Experimenten.

Auch wenn in diesen Experimenten kein Hinweis auf eine explizite Sensitivität für strukturelle Implikationen gefunden wurde, ist eine solche möglicherweise mittels eines anderen Paradigmas nachweisbar. So wurde von anderen Forschern in Bezug auf kausale Ketten eine explizite Sensitivität berichtet (Ahn & Dennis, 2000). Wie in den oben vorgestellten Experimenten erlernten Probanden in diesen Experimenten unabhängig voneinander zwei Kausalzusammenhänge, die ein kausale Kette bildeten. Daraufhin sollten

die Probanden den Zusammenhang zwischen der ersten Ursache und dem letzten Effekt einschätzen. Dabei wurde direkt nach der Stärke des Zusammenhangs gefragt. Der Befund war, dass über eine Vielzahl verschiedener Zusammenhangsstärken hinweg, die von den Probanden geschätzten Zusammenhänge den normativ implizierten sehr nahe kamen. Diese explizite Sensitivität ist möglicherweise auf die spezielle Kausalstruktur zurückzuführen. Bei einer kausalen Kette besteht zwischen der ersten Ursache und dem letzten Effekt ein vermittelter kausaler Zusammenhang. Bei einem Gemeinsame-Ursache und einem Gemeinsamer-Effekt Modell sind die beiden Effekte bzw. Ursachen aber nur über einen implizierten und nicht über einen kausalen Zusammenhang verbunden. Möglicherweise ist eine explizite Sensitivität für strukturelle Implikationen auf vermittelte Kausalzusammenhänge beschränkt. Zur Klärung dieser Frage ist weitere Forschung dringend geboten.

Was bedeuten diese Ergebnisse für unsere Frage nach der Sensitivität für kausale Strukturen?

Die Ergebnisse zeigen, dass Probanden auch sensitiv für die strukturellen Implikationen verschiedener Kausalmodelle sind. Diese Sensitivität ist aber wahrscheinlich nicht auf ein explizites Verständnis der unterschiedlichen Implikationen zurückzuführen. Die Probanden verfügen vielmehr mit der mentalen Simulation über eine Heuristik, welche die Sensitivität gleichsam automatisch eingebaut hat. Indem ein bestimmtes Kausalmodell simuliert wird, ergeben sich die Implikationen von sich aus.

Hinweise auf ein explizites Wissen über die Implikationen verschiedener Strukturen für den Zusammenhang zwischen nicht direkt kausal verbundenen Ereignissen ergaben sich keine. D.h., die Versuchsteilnehmer schienen nicht bewusst zu erwarten, dass die Effekte einer gemeinsamen Ursache in allen Fällen kovariieren, und dass die Ursachen eines gemeinsamen Effektes nicht miteinander zusammenhängen.

Wie passen diese Ergebnisse zu den Befunden aus der Befragungsstudie? Bei dieser hatte sich gezeigt, dass die Probanden bei einem Gemeinsame-Ursache Modell Fälle erwarten, in denen mit der Ursache auch beide Effekte auftreten. Dagegen hatten bei einem Gemeinsamer-Effekt Modell die befragten Probanden erwartet, dass bei Vorliegen des Effektes entweder eine oder beide Ursachen gegeben sind. Beide Erwartungen spiegeln sich in den Einzelvorhersagen der letzten drei Experimente wieder. Es erscheint plausibel anzunehmen, dass die Probanden in der Befragungsstudie ebenfalls durch mentale Simulation zu ihren Urteilen gekommen sind. Das würde auch erklären, weshalb sich die dort geäußerten Erwartungen nicht in den expliziten Urteilen in den Experimenten 6 bis 8 gezeigt haben.

Das Fehlen einer expliziten Sensitivität für strukturelle Implikationen stimmt auch mit den Befunden im ersten Experiment überein. Bei diesem war den Probanden nicht aufgefallen war, dass die gezeigten Daten mit keinem der vorgelegten Strukturmodelle vereinbar waren. Die Experimente zum Prüfen kausaler Strukturhypothesen hatten ebenfalls keine Hinweise auf eine explizite Sensitivität erbracht. Keiner der Probanden in diesen Experimenten hatte die strukturellen Implikationen der Modelle überprüft.

4. DISKUSSION

Zu Beginn dieser Arbeit wurden zwei Hauptfragen gestellt. Einerseits die theoretische Frage, worin sich verschiedene Kausalmodelle unterscheiden und andererseits die empirische Frage, ob Probanden für diese Unterschiede sensitiv sind. Auf beide Fragen wurden in den verschiedenen Kapiteln erste Antworten gegeben. Ziel des ersten Teils dieses abschließenden Kapitels ist es, nochmals die wichtigsten Punkte zusammenzufassen. Dabei wird die Vielzahl der empirischen Ergebnisse auf drei Hauptbefunde verdichtet werden. Zwei theoretische Gesichtspunkte sollen dann ins Zentrum der Aufmerksamkeit gerückt werden: Die Unterscheidung zwischen einer expliziten und einer impliziten Sensitivität für strukturelle Implikationen und die Unterscheidung zwischen einem Denken mit und einem Denken über Kausalmodelle. Eine Analogie aus dem Bereich des physikalischen Denkens soll diesen Unterschied weiter veranschaulichen. Der zweite Teil dieses Kapitels bringt eine neue Sichtweise auf kausales Denken ins Spiel. Anstelle der normativen Sicht, an der wir uns im bisherigen Verlauf dieser Arbeit orientiert hatten, werden wir uns nun einer evolutionären Sicht zuwenden. Die Grundannahme dieser Sicht ist, dass sich unser Lernen und Denken im Laufe der Evolution an die Anforderungen unserer Umwelt angepasst hat und für unser Überleben von Vorteil ist. Den Anforderungen unserer natürlichen Umwelt werden dann die Anforderungen der Wissenschaft an unser Denken gegenübergestellt. Die Unterschiede zwischen den beiden Anforderungen spiegeln sich auch in den Befunden der durchgeführten Experimente wieder. Den Abschluss werden daher Spekulationen über einen möglichen evolutionären Hintergrund der aufgezeigten Differenzen zwischen einem Denken mit und einem Denken über Kausalmodelle bilden.

4.	DISKUSSION	
4.1	Zusammenfassung der Befunde	4.2 Evolutionspsychologische Sichtweise kausalen Denkens
4.1.1	Implizite vs. explizite Sensitivität für strukturelle Implikationen	4.2.1 Lernen durch Beobachten
4.1.2	Denken mit und über Kausalmodelle	4.2.2 Lernen durch Handeln
4.1.3	Analogie zwischen physikalischem und kausalem Denken	4.2.3 Anforderungen der Wissenschaft an kausales Denken
		4.2.4 Denken mit und über Kausalmodelle - Reprise

Abb. 4.1: Struktur Kapitel 4

4.1 Zusammenfassung der Befunde

Die erste zentrale Grundfrage, die am Anfang dieser Arbeit stand, war, welche Unterschiede es zwischen verschiedenen Kausalmodellen gibt. Diese Fragestellung wurde in den ersten beiden Kapiteln untersucht. Dabei stellte sich die Struktur des Modells als die entscheidende Eigenschaft heraus. Die Struktur bestimmt, welche statistischen Indikatoren die Stärke der beteiligten Zusammenhänge angeben, und sie impliziert bestimmte Abhängigkeiten zwischen den nicht direkt kausal verbundenen Ereignissen. Es gibt drei grundlegende Strukturen, auf die alle komplexen Kausalmodelle zurückgeführt werden können. Bei diesen handelt es sich um das Gemeinsame-Ursache Modell, das Gemeinsamer-Effekt Modell und das kausale Kettenmodell. Die beiden Erstgenannten wurden als die beiden wichtigsten Modelle in dieser Arbeit in neun Studien untersucht. Das Gemeinsame-Ursache Modell mit einer Ursache und zwei zugehörigen Effekten impliziert, dass die unbedingte Kontingenz bzw. die unbedingte Power (vgl. Kapitel 1.3) zwischen der Ursache und jedem der beiden Effekte indikativ für die Stärke des Kausalzusammenhangs ist. Darüber hinaus impliziert die Struktur, dass die beiden Effekte miteinander kovariieren sollten. Dieser Zusammenhang sollte aber verschwinden, wenn die Ursache konstant gehalten wird, d.h. die beiden Effekte sollten konditional auf die Ursache unabhängig sein. Das Gemeinsamer-Effekt Modell hat Implikationen, die denen des Gemeinsame-Ursache Modells diametral entgegengesetzt sind. Beeinflussen zwei Ursachen einen gemeinsamen Effekt, dann muss der Einfluss der konfundierenden anderen Ursache kontrolliert werden. In diesem Fall ist die bedingte Kontingenz bzw. die bedingte Power ein valider Indikator für die Stärke der Kausalrelation. Die Gemeinsamer-Effekt Struktur impliziert keinen Zusammenhang zwischen den beiden Ursachen. Sie sollten daher unabhängig voneinander sein. Bedingt auf das Vorliegen des Effektes sollten die beiden Ursachen aber negativ abhängig werden. Diese unterschiedlichen Implikationen der verschiedenen Modelle bilden die Grundlage für alle Anwendungen von Kausalmodellen.

Die zweite Hauptfrage war, ob Probanden für die Unterschiede zwischen den Kausalmodellen sensitiv sind. Zur Beantwortung dieser Frage wurde im zweiten Kapitel eine Fülle von bisher veröffentlichten Befunden erörtert. Dabei zeigte sich, dass sich die vorliegenden Arbeiten auf das Lernen von Kausalzusammenhängen bzw. die Bestimmung der Stärke des kausalen Einflusses konzentriert haben (vgl. 2.1). Die übereinstimmenden Befunde waren, dass Probanden hierbei zwischen Gemeinsame-Ursache und Gemeinsamer-Effekt Modellen unterschieden und diese bei der Ermittlung des Kausalzusammenhangs normativ richtig nutzten.

Als eine weitere erfolgreiche Anwendung von Kausalmodellen erwiesen sich Diagnosen und Prognosen (siehe 2.2). Allerdings haben sich in diesem Bereich nur wenige Studien direkt mit unterschiedlichen Kausalmodellen beschäftigt. Dennoch sprechen die Befunde dafür, dass auch hier die Implikationen der Modelle Beachtung finden. So ließen sich die Urteile der Probanden über ein Bayesianisches Kausalmodell vorhersagen.

Eine insbesondere für die Wissenschaft wichtige Anwendung von Kausalmodellen ist die Prüfung von kausalen Strukturhypothesen (vgl. 2.3.3). Eine solche Prüfung ist unter gewissen Bedingungen auch anhand von Beobachtungsdaten möglich. Das einzige valide Kriterium bei einer solchen Prüfung sind die strukturellen Implikationen. Nur wenn die vermuteten Kausalzusammenhänge vorliegen und die von diesen und der Struktur implizierten Zusammenhänge zwischen den nicht kausal verbundenen Ereignissen gegeben sind, kann die Strukturhypothese zutreffen. Empirische Untersuchungen zu dieser Anwendung von Kausalmodellen lagen vor dieser Arbeit noch nicht vor.

Eine wichtige Rolle spielen die strukturellen Implikationen auch für die Generierung von Kausalmodellen (vgl. 2.4). Ein gut gewähltes Kausalmodell vermag eine Vielzahl von beobachteten Abhängigkeiten und Unabhängigkeiten durch eine begrenzte Zahl von Kausalzusammenhängen zu erklären. Daher kann aus dem Abhängigkeitsmuster der Ereignisse auch das zugrundeliegende Kausalmodell rekonstruiert werden. Dass dies technisch möglich ist, konnte bereits gezeigt werden (vgl. Glymour & Cooper, 1999; Pearl, 2000). Dass Probanden dies können, erscheint aber sehr unwahrscheinlich. Ein psychologisch plausibles Vorgehen bei der Generierung von kausalen Strukturhypothesen ist die Integration von separat erlernten Kausalzusammenhängen (vgl. 2.4.2). Eine erfolgreiche Integration sollte damit einhergehen, dass aus dem entstandenen Modell die entsprechenden strukturellen Implikationen abgeleitet werden können. Auch hierzu gab es bisher keine einschlägigen Studien.

Im dritten Kapitel wurden die im Rahmen dieser Arbeit durchgeführten empirischen Untersuchungen vorgestellt. Die Fragestellung des ersten Experimentes war, ob Probanden bei allen grundlegenden Kausalmodellen die normativ richtigen Indikatoren für die Bestimmung des Kausalzusammenhangs heranziehen (siehe 3.1). Damit erweitert dieses Experiment die vorliegenden Befunde auf kausale Kettenmodelle. Die Ergebnisse sprechen dafür, dass bei allen Modellen normativ richtig vorgegangen wird. Beim Gemeinsamer-Effekt Modell wird der Einfluss der zweiten Ursache kontrolliert und die bedingte Kontingenz bzw. bedingte Power als Maß für die Stärke des Kausalzusammenhangs verwendet. Bei allen anderen Modellen wird die unbedingte Kontingenz bzw. Power als Maß herangezogen, was ebenfalls korrekt ist. Zusätzlich konnte belegt werden, dass dieser Effekt der Kausalstruktur nicht vom Inhalt des zu untersuchenden Sachverhaltes abhängig ist.

Die Fragestellung der Experimente 2-5 war, ob Personen bei der Prüfung kausaler Strukturhypothesen die strukturellen Implikationen der jeweiligen Modelle als Entscheidungskriterium nutzen (siehe Kapitel 3.2). Hierbei wurden das Gemeinsame-Ursache und das Gemeinsamer-Effekt Modell mit ihren entgegengesetzten strukturellen Implikationen miteinander verglichen. Die Befunde aus den vier durchgeführten Experimenten belegen, dass Probanden die strukturellen Implikationen bei der Prüfung nicht berücksichtigen. Die Ergebnisse sprechen vielmehr dafür, dass die Versuchsteilnehmer lediglich das Vorliegen der vermuteten Kausalzusammenhänge überprüfen. Ein solches Vorgehen vermag nicht zwischen verschiedenen Modellen, in denen dieselben Ereignisse kausal verbunden sind, zu differenzieren. Eine Abweichung von diesem Vorgehen ergab sich lediglich dann, wenn das Auftreten eines der Ereignisse für die anderen notwendig war. In diesem Fall wurde dieses Ereignis richtig als gemeinsame Ursache erkannt.

Die Fragestellung der Experimente 6-8 war, ob Probanden in der Lage sind, zwei separat gelernte Kausalzusammenhänge in ein einheitliches Modell zu integrieren und aus diesem Modell Vorhersagen über den Zusammenhang zwischen den beiden nicht direkt kausal verbundenen Ereignissen abzuleiten (siehe 3.3). Verglichen wurden dabei erneut das Gemeinsame-Ursache Modell, welches einen Zusammenhang zwischen den beiden Effekten impliziert, und das Gemeinsamer-Effekt Modell, welches eine Unabhängigkeit zwischen den beiden Ursachen annimmt. Die Befunde belegen, dass Probanden eine solche Integration durchführen können und sie zumindest implizit für die strukturellen Implikationen der entstandenen Modelle sensitiv sind. Die implizite Sensitivität zeigte sich über eine Reihe von Vorhersagen für konkrete Einzelfälle hinweg. So generierten die Probanden bei einem Gemeinsame-Ursache Modell einen Zusammenhang zwischen den Effekten und bei einem Gemeinsamer-Effekt Modell eine Unabhängigkeit zwischen den Ursachen. Bei der direkten Frage nach dem erwarteten Zusammenhang zwischen den Ursachen bzw. den Effekten differenzierten die Probanden aber nicht zwischen den verschiedenen Modellen. Es konnte also keine explizite Sensitivität für strukturelle Implikationen nachgewiesen werden. Der Unterschied zwischen dem expliziten und dem impliziten Maß ist vermutlich darauf zurückzuführen, dass die Probanden bei den Einzelvorhersagen die Kausalmodelle mental simuliert haben. Eine solche Simulation führt automatisch zu einer Sensitivität für strukturelle Implikationen, ohne dass ein explizites Wissen über diese notwendig wäre. Die Antwort auf die explizite Frage haben sie vermutlich geraten.

Fassen wir die Ergebnisse zusammen, dann ergeben sich folgende drei Hauptbefunde. Erstens unterscheiden Probanden beim Lernen von Kausalzusammenhängen bzw. bei der Bestimmung der Stärke des Kausaleinflusses zwischen verschiedenen Kausalmodellen und verwenden je nach Modell die normativ richtigen statistischen Indikatoren. Zweitens ziehen Probanden bei der Prüfung kausaler

Strukturhypothesen nicht die strukturellen Implikationen als Kriterium heran, obwohl diese das einzige valide Kriterium sind. Für eine Widerlegung der strukturellen Implikationen durch die gezeigten Daten sind sie nicht sensitiv. Der dritte Hauptbefund ist, dass Probanden bei Einzelvorhersagen sensitiv für die strukturellen Implikationen sind, sie aber gleichzeitig kein explizites Verständnis für diese zeigen.

Die zweite Hauptfrage, ob Probanden für die Unterschiede zwischen verschiedenen Kausalmodellen sensitiv sind, muss also differenziert beantwortet werden. Grundsätzlich kann die Frage bejaht werden. Alle Experimente zeigen, dass die Probanden Unterschiede zwischen verschiedenen Kausalmodellen machen. Jedoch waren dies nicht immer die normativ richtigen. Normativ gingen die Versuchsteilnehmer nur bei der Bestimmung der Stärke eines Kausalzusammenhangs vor. Sie taten dies aber weder bei der Prüfung von Strukturhypothesen, noch bei der expliziten Einschätzung des Zusammenhangs nicht direkt kausal verbundener Ereignisse.

4.1.1 Implizite vs. explizite Sensitivität für strukturelle Implikationen

Ein wichtiger Teilaspekt der zweiten Hauptfrage ist, ob Personen für die strukturellen Implikationen der Modelle sensitiv sind. Hier zeigte sich in den letzten Experimenten eine interessante Dissoziation. Die Ergebnisse belegen eine implizite Sensitivität für strukturelle Implikationen. Bei den Prognosen bzw. Diagnosen über eine Reihe von Einzelfällen hinweg wurde von den Versuchsteilnehmern die von dem jeweiligen Modell implizierte Abhängigkeit bzw. Unabhängigkeit der nicht direkt kausal verbundenen Ereignisse erzeugt. Im Gegensatz dazu zeigen die Antworten auf die direkte Frage nach dem implizierten Zusammenhang, dass keine explizite Sensitivität vorliegt.⁵⁸

Wie kommt es zu dieser Dissoziation? Eine mögliche Erklärung ist, dass die gefundene implizite Sensitivität nicht die Folge eines bewussten oder unbewussten Verständnisses für strukturelle Implikationen, sondern eine Folge der eingesetzten Strategie bei der Lösung der jeweiligen Aufgabe war. Wenn das Vorgehen selbst sensitiv für strukturelle Implikationen war, dann wiesen die Antworten auf eine Sensitivität hin, ansonsten nicht. Bei den Einzelvorhersagen wurde als Strategie vermutlich auf mentale Simulation zurückgegriffen, ein Prozess, der per se sensitiv für die Implikationen des angenommenen Modells ist. Bei der direkten Einschätzung derselben Zusammenhänge wurde aber (sofern möglich) die gesehene Kovariation der Ereignisse herangezogen. Ein

⁵⁸ Explizit und implizit sind hier nicht synonym zu bewusst und unbewusst. Eine implizite Sensitivität bedeutet nicht, dass sich die Probanden ihrer Sensitivität nicht bewusst sind. Es bedeutet, dass sich im Verhalten aber nicht in den Aussagen der Versuchsteilnehmer eine Berücksichtigung struktureller Implikationen zeigt.

solches Vorgehen ignoriert strukturelle Implikationen. Dasselbe zeigte sich bei der Prüfung von Strukturhypothesen. Hierbei wurden von den Probanden einzelne Kausalzusammenhänge auf ihr Vorhandensein überprüft, ein Vorgehen, das ebenfalls nicht sensitiv für strukturelle Implikationen ist.

Wenn aber die gefundene implizite Sensitivität darauf zurückzuführen ist, dass das gewählte Vorgehen sensitiv ist, dann sind die Befunde mit der Interpretation vereinbar, dass Personen kein Wissen über die strukturellen Implikationen haben. Sie verfügen aber über Strategien wie die mentale Simulation, welche eine Sensitivität von sich aus beinhalten.

4.1.2 Denken mit und Denken über Kausalmodelle

Worin liegen die Unterschiede zwischen den Aufgaben, in denen die Probanden normativ korrekt vorgingen und solchen, in denen ihr Vorgehen zu erheblichen Abweichungen von den Vorhersagen normativer Modelle führte? Meines Erachtens liegt der Hauptunterschied darin, dass ein Teil der Aufgaben ein Denken mit Kausalmodellen erfordert, während die anderen Aufgaben ein Denken über Kausalmodelle verlangen. Beim Denken mit Kausalmodellen muss ein Kausalmodell lediglich angewandt werden. Die Gültigkeit des Modells selbst wird dabei nicht in Frage gestellt. Ein paradigmatischer Fall hierfür sind Aufgaben zum Kausallernen, bei denen ein Kausalmodell vorgegeben wird und die Probanden gebeten werden, die Stärke des Kausaleinflusses der Ursachen einzuschätzen. Beim Denken über Kausalmodelle wird die Gültigkeit des Modells selbst in Frage gestellt. Paradigmatisch hierfür sind die Aufgaben zur Prüfung kausaler Strukturhypothesen.

Wenden wir diese Unterscheidung auf die vorliegenden Befunde an, so zeigt sich folgendes Bild. Die Personen haben eine durchweg hohe Kompetenz beim Denken mit Kausalmodellen. Sehr gute, extrem akkurate Ergebnisse erzielen sie bei der Einschätzung der Effektivität des eigenen instrumentellen Handelns. Ebenfalls gute Ergebnisse zeigen sich bei der Bestimmung der Stärke eines Kausalzusammenhangs innerhalb eines Kausalmodells und bei der Prüfung einfacher kausaler Zusammenhangshypothesen. Ebenso wird bei Diagnosen und Prognosen in Einzelfällen das Kausalmodell lediglich angewandt. Auch hier sprechen die im zweiten Kapitel zitierten Befunde für eine hohe Kompetenz. Bei mentalen Simulationen wird ebenfalls ein Kausalmodell angewandt. Diese führen, wie bereits mehrfach gesagt, zu einer Sensitivität für strukturelle Implikationen.

Beim Denken über Kausalmodelle zeigt sich im Gegensatz hierzu eine eher geringe Kompetenz. Die Leistung der Probanden in den Experimenten zum Hypothesenprüfen war gering. Sie waren nicht in der Lage, systematisch das normativ korrekte Modell herauszufinden. Das Vorgehen der Probanden war zudem ungeeignet für eine Prüfung von Strukturhypothesen. Die Vorhersage des Zusammenhangs zwischen Ereignissen, die innerhalb eines Modells nicht direkt kausal verbunden sind, verlangt ebenfalls ein Denken über Kausalmodelle. Die strukturellen Implikationen müssen entweder bekannt sein oder sie müssen zur Beantwortung der Frage aus dem Kausalmodell erschlossen werden. In keinem der letzten drei Experimente waren die Probanden dazu in der Lage. Neben diesen Ergebnissen aus der experimentalpsychologischen Forschung sprechen auch die Erfahrungen aus der Methodenlehreausbildung für eine hohe Schwierigkeit des Denkens über Kausalmodelle (Waldmann, persönliche Mitteilung).

Bei diesem Überblick sind wir zugebenermaßen über eine Fülle von Detailbefunden hinweggegangen. Dennoch trifft dieses grobe, holzschnittartige Bild meines Erachtens einen wesentlichen Punkt. Die Trennungslinie bei der Kompetenz im kausalen Denken verläuft zwischen einem Denken mit und einem Denken über Kausalmodelle. Dafür spricht auch, dass die eingesetzten Heuristiken beim Denken mit Kausalmodellen sehr erfolgreich sind und nur zu unwesentlich schlechteren Leistungen führen als das normativ richtige Vorgehen. Dagegen weichen die Ergebnisse der Strategien beim Denken über Kausalmodelle erheblich von den normativen Vorhersagen ab.

4.1.3 Analogie zwischen physikalischem Denken und kausalem Denken

Beim physikalischen Denken zeigen sich Phänomene, die eine Analogie zu den obigen Befunden zum kausalen Denken erlauben. Dies gilt insbesondere für die Physik der Wurfbewegungen. Werden Erwachsene und Kinder gefragt, welchen Weg ein Gegenstand nimmt, der von einem bewegten Träger (z.B. einem Flugzeug) fällt, so glauben auch viele Erwachsene, dass der Gegenstand gerade zu Boden fallen wird (Kaiser et al., 1985). Tatsächlich wird der Gegenstand aber eine Parabel beschreiben. Das explizite Urteil ist also falsch. In einem anderen Experiment wurden die Versuchsteilnehmer gebeten, eine Kugel so über einen Tisch zu rollen, dass sie in einer bestimmten Entfernung zu Boden fällt. Das Ergebnis war, dass selbst Kindergartenkinder diese Aufgabe sehr gut lösen können (Krist et al., 1993). Das implizite Urteil, das sich in der Abwurfgeschwindigkeit ausdrückte, war nahezu perfekt. Im gleichen Experiment wurden andere Probanden gebeten, die Kugel nicht zu rollen, sondern die Abwurfgeschwindigkeit explizit zu schätzen. Aus einer Vielzahl von

Urteilen über verschiedene Tischhöhen und Zielentfernungen wurden dann die Urteilsstrategien rekonstruiert. Das Ergebnis war, dass teilweise selbst Erwachsene die für die Bestimmung der Geschwindigkeit relevanten Faktoren nicht richtig integrierten und deshalb nicht zu durchgängig richtigen Urteilen kamen (Krist et al., 1993; vgl. auch Wilkening & Krist, 1995).

Wie in den Befunden zum kausalen Denken zeigt sich beim physikalischen Denken eine hohe Kompetenz nur in den impliziten, aber nicht in den expliziten Maßen. Die Aufgabe, selbst die Kugel zu rollen, entspricht der Anwendung eines Modells über Wurfbewegungen. Sie könnte beispielsweise durch eine mentale Simulation der Bewegung gelöst werden. Die Aufgabe, die Abwurfgeschwindigkeit anzugeben, erfordert es, die angegebenen Informationen über die richtige physikalische Gesetzmäßigkeit miteinander zu verbinden. Dabei muss über die Gesetze der Wurfbewegung nachgedacht werden. Dies entspricht einem Denken über Kausalmodelle. Die Ergebnisse belegen, dass uns dies erheblich schwerer fällt. Sowohl beim physikalischen wie beim kausalen Denken kommt es also zu einer Dissoziation zwischen den expliziten Urteilen und anderen Maßen.

Wenn diese Analogie passend ist, können wir vielleicht aus ihr weitere Schlüsse ziehen. Was die Befunde für das physikalische Denken zeigen, ist, dass für ein erfolgreiches Handeln in der Umwelt ein explizites Verständnis aller ihrer Eigenschaften nicht notwendig ist. Möglicherweise trifft dies auch auf kausales Denken zu. Für eine erfolgreiche Anwendung von Kausalmodellen beim Lernen, beim Prognostizieren und Diagnostizieren wäre danach ein explizites Verständnis der strukturellen Implikationen dieser Modelle nicht erforderlich. Diese Vermutung würde gut zu der oben dargestellten Analyse passen, dass die gefundene implizite Sensitivität auf die Eigenschaften der eingesetzten Strategien und nicht auf Wissen zurückzuführen ist (vgl. 4.1.2).

Wie bei jeder Analogie gibt es auch Punkte, in denen die beiden analogen Phänomene voneinander abweichen. Beim physikalischen Denken unterscheidet sich das gezeigte Verhalten von den gefälltten Urteilen. Dagegen unterscheiden sich beim kausalen Denken die Urteile in verschiedenen Aufgaben, welche dieselbe Fragestellung betreffen. Trotz dieser Abweichung ist der Bereich des physikalischen Denkens meines Erachtens ein anschauliches Beispiel dafür, wie sich die Anwendung eines Modells (Werfen \leftrightarrow Kausallernen) von dem Denken über ein Modell (Bestimmung physikalischer Gesetze \leftrightarrow Prüfen von Kausalmodellen) unterscheidet.

4.2 Evolutionspsychologische Sichtweise kausalen Denkens

Bereits im ersten Kapitel haben wir die Bedeutung von Kausalmodellen umrissen. Durch ein solches Modell können wir einen Sachverhalt und die zugehörigen kausalen Mechanismen verstehen. Sie ermöglichen es uns, zielgerichtet zu handeln und gewünschte Effekte hervorzurufen. Phänomene können mit ihrer Hilfe erklärt werden und sie erlauben es zukünftige Ereignisse vorherzusagen. Betrachten wir diese Punkte, so fällt auf, dass dies Merkmale sind, die direkt für unser Überleben als biologische Organismen eine Bedeutung haben. Die Deutung von Spuren als eine Form der Diagnose ist wichtig für das Vermeiden von Feinden und das Finden gesuchter Nahrungsmittel. Die Prognose von Ereignissen ist wichtig dafür, aversive oder gar lebensbedrohliche Ereignisse zu vermeiden und erwünschte Effekte abzuwarten. Das Verstehen beispielsweise natürlicher Wachstumsprozesse erlaubt das Betreiben von Ackerbau. Auf der Grundlage von Kausalmodellen können Handlungen geplant und Effekte dieser mental vorweggenommen werden. Das mentale Durchspielen einer Handlungsoption erspart dabei ein direktes Handeln und schont so wichtige Ressourcen. Auch können Handlungen durch das Wissen um wichtige Einflussfaktoren optimal an die vorhandenen Bedingungen angepasst werden. Angesichts dieser Überlegungen scheint die Behauptung gerechtfertigt, dass unsere Fähigkeit, kausale Modelle über Sachverhalte in der Welt zu entwickeln, uns einen wichtigen Überlebensvorteil sichert.

Ziel dieses abschließenden Teilkapitels ist es daher, weg von einer normativen Sichtweise von Kausalmodellen hin zu einer evolutionspsychologischen Sichtweise zu gehen. Die Grundidee der Evolutionspsychologie ist, dass sich unser Lernen und Denken im Laufe der Menschheitsgeschichte an die natürlichen Umgebungsbedingungen angepasst hat. Anstelle der Frage, welche Eigenschaften und normativen Implikationen Kausalmodelle haben, soll die Frage daher nun lauten, welche Anforderungen unsere Umwelt an uns stellt und wie diese Anforderungen unser Denken geformt haben. Eine erste Antwort auf diese Frage ergibt sich dabei direkt aus den obigen Ausführungen. Die grundlegende Anforderung an unser Lernen und Denken ist, die Kausalzusammenhänge, die für unser Überleben entscheidend sind, zu erkennen und dieses Wissen für unser Handeln nutzbar zu machen.⁵⁹

Für den Erwerb von Kausalwissen stehen uns zwei Wege offen. Einerseits können wir unsere Umwelt beobachten, andererseits können wir aktiv handelnd in diese eingreifen. Beide Wege sind für uns von Bedeutung, da wir nicht alle relevanten Faktoren direkt beeinflussen können. Zum Beispiel können wir, um herauszufinden, ob durch eine

⁵⁹ Eine Alternative zu wissensbasiertem Handeln ist eine genetisch verankerte Verhaltensadaptation. Das durch eine solche Adaptation resultierende Verhalten ist aber häufig wenig flexibel und kann nicht kurzfristig an Veränderungen in der Umwelt angepasst werden. Wissensbasiertes Handeln ist daher genetisch basierendem Verhalten in vielen Fällen überlegen.

Beseitigung anderer Gräser der Ertrag an Getreide gesteigert werden kann, einen Teil des Ackers jäten. Um herauszufinden, welchen Einfluss das Wetter auf den Ertrag hat, können wir nur beobachten. Beides ist wichtig, wenn es darum geht, sich auf eine eventuelle Nahrungsmittelknappheit im Winter vorzubereiten.

Der erste Abschnitt des Teilkapitels wird sich mit Lernen durch Beobachtung beschäftigen. Unsere Umwelt bietet sich uns als ein kontinuierlicher Strom von Ereignissen dar, der stets von der Ursache zum Effekt verläuft. Es ist davon auszugehen, dass sich unser kausales Denken an diese Bedingungen angepasst hat. Der zweite Abschnitt geht auf das Lernen durch Handlungen ein. Die meisten für uns relevanten Kausalrelationen können wir direkt durch unser Handeln beeinflussen und damit Wissen über diese erwerben. Im dritten Abschnitt werden wir uns den Anforderungen zuwenden, welche die moderne Naturwissenschaft an unser kausales Denken stellt, und werden diese mit den Anforderungen unserer natürlichen Umwelt vergleichen. Im letzten Teilkapitel werden die in dieser Arbeit verwendeten Aufgabenstellungen und die empirischen Befunde aus der Sicht der Evolutionspsychologie diskutiert.

4.2.1 Lernen durch Beobachten

Durch Beobachten kann Wissen erworben werden sowohl über Kausalzusammenhänge, die wir selbst beeinflussen können, als auch über solche, auf die wir keinen Einfluss haben. Beide Arten von Kausalzusammenhängen sind für uns von Bedeutung. Zusätzlich können Zusammenhänge zwischen Ereignissen erlernt werden, die zwar nicht selbst kausal verbunden sind, die aber auf anderen Kausalzusammenhängen beruhen. So ist es beispielsweise wichtig zu erkennen, wann der optimale Zeitpunkt für eine Aussaat ist, wohin das Wild im Winter zieht und welche Anzeichen ein Unwetter ankündigen. Obwohl dies keine direkten Kausalzusammenhänge sind, sind diese Zusammenhänge eine Folge anderer Kausalrelationen. Nehmen wir zum Beispiel die Bestimmung des richtigen Zeitpunkts für die Aussaat. Die hierfür kausal relevante Variable ist die Temperatur. Sinken die Temperaturen nach der Aussaat erheblich, dann ist die Ernte gefährdet. Die Temperaturen hängen aber stark von der Jahreszeit ab. Diese wiederum lässt sich eindeutig mit bestimmten Sternkonstellationen in Verbindung bringen. Die Erkenntnis, dass durch die Beobachtung des Sternhimmels der optimale Zeitpunkt für die Aussaat bestimmt werden kann, war eine sehr wichtige Erkenntnis für die Menschheit. Zwar hat der Sternhimmel entgegen der Behauptung der Astrologen keinen direkten kausalen Einfluss auf das Leben auf der Erde, aber die Sternkonstellation ist genauso wie die Jahreszeiten eine Folge der Bahn der Erde um die Sonne und der Ekliptik. Wir haben es also mit einem

Gemeinsame-Ursache Modell zu tun, wobei die Stellung der Erde zur Sonne die gemeinsame Ursache für den beobachteten Stand der Gestirne und die Jahreszeit ist. Bei der Vorhersage des Aussaatzeitpunktes wird die von diesem Modell implizierte Abhängigkeit zwischen Temperatur und Sternkonstellation genutzt.

Eine Beobachtung der Ereignisse in der Welt resultiert also entweder in einem Wissen über direkte Kausalzusammenhänge oder in einem Wissen über Zusammenhänge, die auf das Wirken anderer Kausalzusammenhänge zurückzuführen sind. Die Anforderung an das Lernen durch Beobachtung ist folglich, diese Zusammenhänge zu erkennen. Doch wie bietet sich uns unsere Umwelt dar? Welche Grundlage haben wir, wenn wir durch Beobachten lernen? Die Welt zeigt sich uns in Form eines unstrukturierten Stroms immer neuer Ereignisse. Dieser Ereignisstrom verläuft dabei stets von der Ursache zum Effekt. Dies ist die Grundlage für unser Lernen.

Der Ereignisstrom der Welt unterscheidet sich erheblich von dem, was in Laborexperimenten zum Lernen den Versuchsteilnehmern (Menschen wie Tieren) dargeboten wird. In Laborsettings gibt es nur eine begrenzte Zahl von Ereignissen und der Strom der Ereignisse ist in klare Trials gegliedert. Dies ermöglicht es dem Lernenden, statistische Zusammenhänge zwischen den Ereignissen direkt zu bestimmen. Das ist in unserer Umwelt so nicht gegeben. Die Strukturierung, welche in Laborexperimenten durch den Versuchsleiter für die Probanden durchgeführt wird, muss unter realen Bedingungen vom Organismus selbst geleistet werden.⁶⁰ Der folgende Abschnitt beschäftigt sich damit, wie dies möglich ist.

Kontinuierlicher Strom von Ereignissen

Unsere Umwelt bietet uns einen kontinuierlichen Strom von Ereignissen dar, die in gewissen Zeitabständen auftreten und wieder verschwinden. Dieser Ereignisstrom weist keine natürlichen Schnitte auf, er wird nur durch das Auftreten der Ereignisse selbst strukturiert. Unter diesen Bedingungen ist ein statistischer Zusammenhang zwischen Ereignissen nicht bestimmbar. Alle Ereignisse treten im Laufe der Zeit nacheinander auf. Für eine Ermittlung von Zusammenhängen ist es zuerst notwendig, die Ereignisse, die zusammengehören, miteinander zu verbinden. Es muss also festgelegt werden, wann sowohl die Ursache als auch der Effekt vorliegen, nur einer von beiden gegeben ist oder beide ausbleiben. Wenn die Häufigkeit dieser Fälle bekannt ist, dann können Kontingenzen oder andere Maße für Kausalität bestimmt werden. Welche Möglichkeiten haben wir zur

⁶⁰ Der aufgezeigte Unterschied zwischen einer Trial-by-Trial Darbietung und einem unstrukturierten Ereignisstrom bedeutet aber nicht, dass Laborstudien einen geringeren Erkenntniswert haben. Diese Analyse zeigt lediglich, dass unter natürlichen Bedingungen vor einer Bestimmung statistischer Zusammenhänge eine Strukturierung des Ereignisstroms stehen muss. Dass Menschen wie Tiere statistische Zusammenhänge als Indikatoren für Kausalität nehmen, steht außer Frage.

Lösung dieser Anforderung? Es sind drei Alternativen denkbar. Eine Alternative ist es, die Zeitintervalle zwischen den Ereignissen als Indikatoren zu nutzen, eine zweite Alternative ist es, ein Zeitintervall a priori festzusetzen, in dem zusammengehörige Ereignisse auftreten müssen. Eine dritte Alternative wäre es, dass Kausalzusammenhänge direkt wahrgenommen werden können. Die drei Alternativen und die entsprechenden empirischen Befunde werden in den folgenden Absätzen vorgestellt.

Eine Möglichkeit für die Bestimmung eines Kausalzusammenhangs ist es, die Zeitintervalle zwischen den Ereignissen zu betrachten. Wenn eine Ursache ein Effektereignis beeinflusst, dann sollten die Zeitintervalle zwischen den Effektereignissen kürzer sein, wenn die Ursache präsent ist, als wenn sie nicht gegeben ist. Ist das Wetter durchgehend warm und feucht, dann ist das Intervall, bis ein Baum erneut Früchte trägt, kurz, häufig nur wenige Wochen. Ist das Wetter kalt, beträgt das Intervall ein Jahr. Das Wetter ist also ein wichtiger Kausalfaktor für das Wachstum von Früchten. Paaren sich Tiere häufig, dann haben sie ständig neuen Nachwuchs, verhindert man die Fortpflanzung, dann kann der Zeitraum bis zum nächsten Wurf beliebig verlängert werden. Die Paarung ist also eine notwendige Voraussetzung für Trächtigkeit. In beiden Fällen ist der Vergleich der Zeitintervalle indikativ für das Vorliegen eines Kausalzusammenhangs. In theoretischen Überlegungen konnte gezeigt werden, dass ein Vergleich von Zeitintervallen zu den gleichen Ergebnissen führt wie eine Bestimmung von Kontingenzen (Gallistel & Gibbon, 2000). Empirische Untersuchungen belegen, dass Versuchstiere eine hohe Sensitivität für Veränderungen der Zeitintervalle haben. Wird bei konstanten Kontingenzen das Zeitintervall zwischen der Ursache und dem Effekt erhöht, aber das Zeitintervall zwischen den Effektereignissen konstant gehalten, dann sinkt die Reaktionsrate (Gallistel & Gibbon, 2000). Dieser Befund spricht eindeutig dafür, dass unter solchen, quasi natürlichen Bedingungen nicht die Kontingenz per se entscheidend für die Bestimmung des kausalen Einflusses ist.⁶¹

Eine zweite Möglichkeit bei einem kontinuierlichen Ereignisstrom Zusammenhänge zu bestimmen ist es, ein kritisches Zeitintervall festzulegen, innerhalb dessen eine Ursache zu einem Effekt führen kann. Tritt die Ursache auf und der Effekt folgt innerhalb des Intervalls nach, dann spricht dies für den Zusammenhang, tritt nur die Ursache oder nur der Effekt allein auf, ohne dass innerhalb des Zeitintervalls das jeweils andere Ereignis aufgetreten ist, so spricht dies gegen einen Zusammenhang. Da die gemeinsame Abwesenheit von Ursache und Effekt der Regelfall ist, spricht dies weder für noch gegen einen Zusammenhang.⁶²

⁶¹ Dieser Befund hat keine Implikationen für die in dieser Arbeit berichteten Ergebnisse. In keiner der durchgeführten Studien wurden den Probanden Informationen über Zeitintervalle zwischen den Ereignissen gegeben. In diesem Fall sind auch nach diesem Ansatz die Kontingenzen entscheidend.

⁶² Dieser Sachverhalt liefert eine Begründung dafür, weshalb das gemeinsame Ausbleiben nicht als positive Evidenz gewertet werden sollte. Interessanterweise geben Probanden diesem Fall selbst in

Von welchem typischen Zeitintervall zwischen Ursache und Wirkung gehen wir als Menschen aus? Für unsere eigenen Handlungen gilt meistens das Kontiguitätsprinzip, d.h., unsere Handlungen führen in der Regel unmittelbar zu ersten Effekten. Es ist daher anzunehmen, dass das Kriterium der räumlichen und zeitlichen Nähe von Ursache und Wirkung auch auf die Beobachtung von Kausalzusammenhängen angewandt wird. In empirischen Untersuchungen zur Phänomenalen Kausalität zeigte sich, dass Ereignisse (Bewegungen von Körpern), die unmittelbar aufeinander folgen, eher als kausal verbunden wahrgenommen werden als Ereignisse, die räumlich und zeitlich getrennt sind (Michotte, 1963). Auch entwicklungspsychologische Untersuchungen sprechen dafür, dass die Kontiguität ein wichtiges Kriterium für die Bestimmung von Kausalität ist (Bullock et al., 1982; Shultz, 1982). Studien zum Umgang mit komplexen Systemen deuten darauf hin, dass Probanden Schwierigkeiten damit haben, Ursache-Wirkungszusammenhänge zu erkennen, die lange Latenzzeiten haben (Dörner, 1989). Ebenso belegen experimentelle Untersuchungen zum instrumentellen Lernen, dass ein Verhalten innerhalb eines bestimmten Zeitraums verstärkt werden muss, damit es zu einer Konditionierung kommt (Domjan, 1998).

Eine dritte Möglichkeit ist, dass wir Kausalzusammenhänge direkt wahrnehmen können. Direkt wahrzunehmen heißt, dass es nicht notwendig ist, mehrfach die Abfolge von Ursache und Wirkung zu beobachten, bevor ein Kausaleindruck entsteht. Evolutionär gesehen, sollte diese Fähigkeit nur auf für uns elementar wichtige Kausalzusammenhänge beschränkt sein. Die meisten unserer Handlungen beruhen auf mechanischen Bewegungen. Diese wären also ein Kandidat. Zu diesen Bewegungen zählen Stossen, Ziehen und Drücken. Die bereits angeführten Studien zu Phänomenaler Kausalität liefern interessante Belege in dieser Richtung. In diesen Untersuchungen wurden den Probanden unterschiedliche Bewegungsmuster von Punkten gezeigt. Den stärksten Kausaleindruck hatten sie, wenn ein bewegter Punkt einen ruhenden Punkt „anstieß“ und dieser sich danach in Bewegung setzte (Michotte, 1963). Ein weiterer Beleg sind Studien mit Säuglingen zur Objektpermanenz. Diese zeigen, dass Objekte bereits kurz nach der Geburt als undurchdringlich angesehen werden (Spelke et al., 1992). Es ist also durchaus möglich, dass wir im Bereich der Mechanik eine angeborene Fähigkeit haben, Kausalzusammenhänge über Beobachtung direkt wahrnehmen zu können, ohne auf statistische Zusammenhänge rekurren zu müssen.

Fassen wir die obigen Befunde zusammen, dann erscheint es so, als ob Menschen alle drei Alternativen anwenden können, um bei einem kontinuierlichen Ereignisstrom Zusammenhänge zwischen Ereignissen zu bestimmen.

Übereinstimmung von zeitlicher und kausaler Abfolge der Ereignisse

Betrachten wir unsere Umwelt, so verlaufen die Kausalzusammenhänge zeitlich immer von der Ursache zur Wirkung. In der Beobachtung entspricht daher die zeitliche Abfolge in der Regel der kausalen Abfolge. Es gibt aber auch seltene Ausnahmefälle, in denen wir beide Ereignisse gleichzeitig sehen oder den Effekt vor seiner Ursache zu Gesicht bekommen. Dies ist zum Beispiel bei Spuren der Fall. Zuerst sehen wir die Fährte eines Tieres und dann das Tier selbst. Dennoch ist die zeitliche Abfolge unter natürlichen Bedingungen ein valides Kriterium für die kausale Abfolge. Wenn sich unser kausales Denken an die Umwelt angepasst hat, dann sollten wir diese Eigenschaft nutzen. Hierfür sprechen sowohl die Studien zur Phänomenalen Kausalität als auch Studien zum instrumentellen Lernen. Eine Verstärkung ist nur dann verhaltensrelevant, wenn sie nach dem instrumentellen Verhalten gegeben wird (Domjan, 1998).

Des Weiteren ist anzunehmen, dass sich unsere Lernmechanismen an die natürliche Abfolge angepasst haben. In der Regel werden zuerst die Ursachen und dann Effekte gesehen. Diese Form des Lernens wird auch als prädiktives Lernen bezeichnet. Eine Vielzahl von Studien hat gezeigt, dass Probanden dies sehr gut beherrschen (vgl. Waldmann, 1994). Assoziationalistische Lernmechanismen (z.B. Rescorla & Wagner, 1972; Shanks et al., 1996) setzen sogar implizit voraus, dass die zeitliche Abfolge der Lernereignisse und die kausale Abfolge der Ereignisse übereinstimmen. In diesen Modellen werden die zeitlich zuerst gesehenen Ereignisse stets als Cues und die darauf folgenden Ereignisse immer als Outcomes verwendet. Stimmen die kausale und die zeitliche Abfolge nicht überein, liefern diese Modelle falsche Vorhersagen (vgl. Kapitel 2.1).

Allerdings werden wir auch in unserer natürlichen Umwelt teilweise mit diagnostischen Problemen konfrontiert, so beim Spurenlesen oder bei medizinischen Diagnosen. Daher sollten wir auch die Fähigkeit haben, in diesen Ausnahmefällen diagnostisch (vom Effekt auf die Ursache) zu lernen. Diagnostisches Lernen stellt aber höhere Anforderungen an die Verarbeitungskapazität als prädiktives Lernen. Beim diagnostischen Lernen muss der Widerspruch zwischen kausaler und gesehener zeitlicher Abfolge beim Aufbau eines Kausalmodells aufgelöst werden. Empirische Untersuchungen belegen, dass wir hierzu in der Lage sind (Waldmann, 2000, in press; Waldmann & Holyoak, 1992). Es gibt aber Hinweise dafür, dass Probanden bei höherer Komplexität nicht mehr in allen Fällen diagnostisch lernen, sondern dazu tendieren, die zuerst gezeigten Effekte so zu behandeln, als ob sie Ursachen wären (Waldmann, 2000). Die Probanden gehen also auf prädiktives Lernen zurück.

Insgesamt deuten die Befunde darauf hin, dass unser kausales Denken und Lernen an die natürliche Abfolge der Ereignisse adaptiert ist. Aber es ist flexibel genug, um auch diagnostische Informationen verarbeiten zu können.

Unterscheidung von Kausalzusammenhängen und nicht-kausalen Zusammenhängen

Lernen durch Beobachtung resultiert in einem Wissen über Zusammenhänge zwischen Ereignissen. Diese Zusammenhänge können entweder direkt kausal sein oder auf anderen, nicht beobachtbaren Kausalzusammenhängen beruhen.⁶³ Die Frage ist, ob wir aufgrund von Beobachtungen zwischen diesen beiden Formen von Zusammenhängen unterscheiden können. Die im Laufe dieser Arbeit besprochenen strukturellen Implikationen von Kausalmodellen würden eine solche Unterscheidung allein auf der Grundlage von Beobachtungsdaten zulassen (vgl. 2.4). Meines Erachtens ist dies evolutionär aber nicht notwendig. Die Unterscheidung zwischen Kausalzusammenhängen und Zusammenhängen, die eine Folge von Kausalzusammenhängen sind, ist nur dann relevant, wenn es um Zusammenhänge geht, die wir durch unser Handeln beeinflussen können. In diesem Fall ist es für uns wichtig zu erkennen, ob wir durch unsere Handlungen den Effekt hervorrufen können. Dieses Wissen können wir aber sehr schnell durch aktives Ausprobieren erwerben. Weshalb sollten wir stattdessen komplexe Bayesianische Berechnungen ausführen? Bei Zusammenhängen, die wir als Menschen nicht direkt beeinflussen können, ist es von untergeordneter Bedeutung, ob es sich um Kausalzusammenhänge handelt oder nicht. Wichtig ist einzig und allein, dass das eine Ereignis zuverlässig das andere Ereignis vorhersagt. Wenn der Stand der Gestirne den richtigen Zeitpunkt für die Aussaat vorhersagt, dann sind die astronomischen Grundlagen hierfür irrelevant.

Fassen wir diese Überlegung zusammen, dann ergibt sich folgendes Bild: Die Beobachtung des Ereignisstroms in der Welt resultiert in einem Wissen über Zusammenhänge zwischen Ereignissen und Hypothesen über Kausalzusammenhänge. Diese Hypothesen können dann durch gezielte Handlungen überprüft werden.

⁶³ In diesem Abschnitt geht es nicht um das Beobachten von Handlungen anderer kausaler Agenten. Eine solche Beobachtung erlaubt, wie das eigene Handeln, natürlich eine Unterscheidung von kausalen und nicht kausalen Zusammenhängen.

4.2.2 Lernen durch Handeln

Das Lernen durch Handeln weist eine Reihe von Parallelen und mehrere wichtige Unterschiede zum Lernen über Beobachtung auf. Gemeinsam ist beiden der unstrukturierte Strom von Ereignissen, den die Welt uns darbietet. Auch bei Handlungen tritt das Problem auf, den zu einer Ursache zugehörigen Effekt zu finden, um damit die Grundlage für die Bestimmung von Kausalzusammenhängen zu schaffen. Die Möglichkeiten zur Lösung dieses Problems haben wir bereits im vorherigen Abschnitt kennengelernt. Diese treffen auch auf den Wissenserwerb durch Handeln zu. Im Unterschied zum Beobachten haben Handlungen aber mehrere entscheidende Vorteile für die Bestimmung von Kausalzusammenhängen. Diese werden in den folgenden Abschnitten vorgestellt.

Vermeiden von Konfundierungen

Im Strom der Ereignisse treten vor einem Effekt viele potentielle Ursachen auf. Dadurch droht ständig die Gefahr einer Konfundierung bei der Bestimmung von Kausalzusammenhängen. Manipulieren wir aber per Zufall die potentielle Ursache, so wird diese Ursache unabhängig von allen anderen Ursachen des Effektes. Damit wird eine Konfundierung ausgeschlossen. Da sich unsere Umwelt ständig wandelt, ergibt sich eine Unabhängigkeit unserer Handlungen von vielen anderen potentiellen Ursachefaktoren häufig von selbst. Der dann ermittelte Zusammenhang gibt an, wie stark wir unter natürlichen Bedingungen durch unser Handeln den Effekt beeinflussen können. Dies ist das für uns als kausale Agenten relevante Wissen.

Durch kausales Handeln unter verschiedenen Bedingungen werden wir sogar herausfinden, bei welchen Gegebenheiten wir besonders gut einen Effekt hervorrufen können. So haben unsere Vorfahren herausgefunden, dass Weinreben besonders gut an windgeschützten Hängen gedeihen, die Jagd in den frühen Morgenstunden erfolgversprechend ist und Wildtiere mit Jungen am aggressivsten sind. Dieses Wissen ist für uns extrem hilfreich und fördert unser Überleben.

Unterscheidung von kausalen und nicht-kausalen Zusammenhängen

Durch die Manipulation von Ereignissen können wir herausfinden, welches die Ursache und welches der Effekt ist (vgl. 2.4). Damit können auch kausale von nicht-kausalen Zusammenhängen unterschieden werden. Wie die Diskussion zu Ende des Abschnitts über Lernen durch Beobachtung gezeigt hat, ist diese Differenzierung sehr wichtig für erfolgreiches Handeln, beispielsweise bei der Behandlung von Kranken. Werden durch das Medikament nur die Symptome gemildert oder führt die Behandlung zu einer Gesundung des Kranken? Beim Lernen durch Handeln wird diese Unterscheidung automatisch gemacht.

Das Lernen über Handlungen führt also zu einem valideren Wissen über Kausalzusammenhänge als Beobachtungen. Es gibt aber noch andere Gründe, die für ein Primat dieser Form des Lernens sprechen. So ist das Wissen darüber, wie wir für uns wünschenswerte Effekte maximieren und für uns negative Folgen minimieren können, direkt überlebensrelevant. Das Lernen von Zusammenhängen, die wir nicht selbst beeinflussen können, steigert zwar auch unsere Überlebenswahrscheinlichkeit, ist aber nicht essentiell. Aus dieser Überlegung folgt erstens, dass Kausallernen über eigene Handlungen eine elementare Fähigkeit sein sollte, über die alle Organismen verfügen. Hierfür sprechen die Erkenntnisse aus der Forschung zum instrumentellen Konditionieren. Diese zeigen, dass Tiere in der Lage sind, ihr Verhalten so an einen Verstärkungsplan anzupassen, dass sie eine maximale Verstärkungsrate erzielen (Domjan, 1998). Zweitens folgt aus obigen Überlegungen, dass Lernen durch Handlungen besonders effektiv sein sollte. Dies bestätigen ebenfalls experimentelle Untersuchungen. Lernen Menschen Kausalzusammenhänge durch eigene Handlungen, so geben ihre Einschätzungen sehr akkurat die tatsächlich vorliegenden Kontingenzen wieder (Shanks, 1993; Shanks et al. 1996; Wasserman et al., 1993). Im Gegensatz dazu berichten Studien, in denen die Probanden lediglich Daten über Zusammenhänge erhalten und selbst nicht handeln dürfen, teilweise erhebliche Abweichungen der Einschätzungen von den tatsächlich vorhandenen Zusammenhängen (Alloy & Tabachnik, 1984; vgl. auch Kapitel 2.1).

Fassen wir die obigen Überlegungen zu einer evolutionspsychologischen Sichtweise zusammen. Die grundlegende Anforderung der Umwelt an unser kausales Lernen und Denken ist, Kausalzusammenhänge richtig zu erfassen und dieses Wissen in erfolgreiches Handeln umzusetzen. Als Grundlage liefert uns die Welt einen kontinuierlichen Strom von Ereignissen, in dem Ursachen von Effekten gefolgt werden. In diese Abfolge können wir handelnd eingreifen und so bestimmte Ereignisse hervorrufen und andere verhindern. An diese Bedingungen haben wir unser kausales Lernen und Denken adaptiert. Durch unser Verhalten erwerben wir auf effiziente Weise ein sehr genaues Wissen über die Folgen unserer Handlungen. Ebenso sind wir in der Lage, mechanische Kausalzusammenhänge sehr schnell allein durch Beobachtung zu erfassen. Andere Zusammenhänge können wir einerseits durch eine Betrachtung der Zeitintervalle zwischen den Ereignissen oder durch eine Analyse der statistischen Zusammenhänge bestimmen. Diese Zusammenhänge sind entweder kausal oder sie beruhen auf anderen Kausalzusammenhängen, für die sie als Indikatoren verwendet werden können. Bei der Bestimmung der Kausalrichtung liefert uns die natürliche Abfolge der Ereignisse ein in der Regel valides Kriterium. Daran haben sich die Lernmechanismen adaptiert. Das über kausales Lernen erworbene Wissen können wir

beim Lernen neuer Kausalzusammenhänge, bei der Planung von Handlungen und der Vorhersage von Ereignissen zu unserem Nutzen einsetzen.

4.2.3 Anforderungen der Wissenschaft an kausales Denken

Die Wissenschaft stellt ebenso wie unsere Umwelt die Anforderung an uns, Kausalzusammenhänge in der Welt zu erkennen. Der Hauptunterschied zwischen beiden liegt vor allem in den jeweiligen Erfolgskriterien. Unser kausales Denken in unserer natürlichen Umwelt ist dann erfolgreich, wenn es uns erlaubt, durch unsere Handlungen für uns als biologische Organismen förderliche Ereignisse vermehrt hervorzurufen. Wissenschaftliches Denken ist dann erfolgreich, wenn es die kausale Struktur eines Sachverhalts richtig erfasst.⁶⁴

Diese unterschiedlichen Erfolgskriterien bringen unterschiedliche Anforderungen mit sich. So erfordert unsere Umwelt nur, dass wir gut an die Bereiche angepasst sind, die für unser Überleben relevant sind. Zu diesen zählen Mechanik, Ernährung und soziale Interaktionen. Die Wissenschaft verlangt dagegen, dass das Denken unabhängig von der jeweiligen Domäne ist.

Für unser Überleben ist es wichtig zu erkennen, wie stark wir unter normalen Bedingungen durch eine Handlung einen Effekt beeinflussen können. Dagegen ist die Wissenschaft daran interessiert, die Stärke der Wirkung per se, unabhängig von spezifischen Randbedingungen zu erfassen. Ein Beispiel soll diesen Unterschied verdeutlichen. Für die Wissenschaft ist es interessant zu wissen, wie stark ein bestimmter Dünger den Ertrag erhöht. Ein Bauer will dagegen wissen, ob sich sein Ertrag steigert, wenn er den Dünger auf seinen Feldern unter den dort vorhandenen Bedingungen einsetzt. Dabei kann es sein, dass ein im Gewächshaus gut wirkender Dünger unter anderen Gegebenheiten keine Wirkung hat, beispielsweise weil der Dünger durch geringe Niederschläge nicht an die Wurzeln der Pflanzen gelangen kann. Für eine Feststellung der Wirkung einer Ursache unter natürlichen Bedingungen reicht einfaches Handeln aus. Dagegen erfordert die Bestimmung der Wirkung unabhängig von spezifischen Bedingungen spezielle Techniken wie Experimente und statistische Methoden. Diese mussten von der Wissenschaft erst neu entwickelt werden.

Ziel der Wissenschaft ist es, die Kausalzusammenhänge, so wie sie in der Welt vorkommen, zu erfassen. Dies erfordert eine Unterscheidung zwischen direkten Kausalzusammenhängen und statistischen Zusammenhängen, die auf anderen

⁶⁴ Diese Korrespondenztheorie der Wahrheit ist eine unter Naturwissenschaftlern weit verbreitete Sichtweise. Sie besagt, dass eine Aussage genau dann wahr ist, wenn der Inhalt der Aussage mit dem Sachverhalt übereinstimmt. Daneben sind auch andere Wahrheitsbegriffe denkbar, z.B. ein kohärenztheoretischer oder pragmatischer Wahrheitsbegriff.

Kausalzusammenhängen beruhen. Dies ist aber, wie oben ausgeführt, nur zum Teil auch eine evolutionäre Anforderung. Wenn es um Handlungen geht, müssen wir zwischen den beiden Arten von Zusammenhängen unterscheiden. Diese Unterscheidung ergibt sich aber bei Ereignissen, die wir selbst beeinflussen können, durch unser Handeln automatisch. Bei Ereignissen, die wir nicht beeinflussen können, ist es für uns häufig nur relevant, dass wir das eine Ereignis durch ein anderes, diesem vorrausgehendes Ereignis vorhersagen können. Für diese Vorhersage ist aber eine genaue Unterscheidung in kausale und nicht-kausale Zusammenhänge nicht notwendig.

Des Weiteren fordert die Wissenschaft, dass die Kausalmodelle alle beobachtbaren Kovarianzen anhand möglichst weniger theoretischer Konstrukte erklären. Die Kausalmodelle sollten also möglichst sparsam sein. Dagegen erfordert ein erfolgreiches Handeln in der Welt, dass sich die Kausalmodelle auf konkret beobachtbare, wenn möglich durch uns beeinflussbare Ereignisse beziehen. Ein gutes Beispiel hierfür bietet die Geschichte der Medizin. Schon in der Antike wurden von Hippokrates mit seiner Theorie der Unausgewogenheit der Körpersäfte theoretische Konstrukte eingebracht. Für die Behandlung von Krankheiten war aber das Wissen über die Wirkung von Arznei- und Nahrungsmitteln sowie über die Auswirkungen, verschiedener Lebensweisen weitaus wichtiger. Der Therapieerfolg hing also weitest gehend von dem Wissen über die Wirkung direkt beeinflussbarer Faktoren ab. Das theoretische Modell war nur insofern von Bedeutung, als auf seiner Grundlage Vorhersagen für die Wirkung neuer Therapieformen abgeleitet werden konnten. Dass die Ableitung theoretischer Konstrukte nicht direkt für unser Überleben relevant ist, belegt auch die Tatsache, dass nur sehr wenige Menschen hierzu überhaupt in der Lage sind.

Daneben gibt es noch weitere Kriterien, die wissenschaftliche Theorien erfüllen müssen. Zu diesen gehören Konsistenz und Kohärenz. Diese Anforderungen gehen weit über das hinaus, was die Natur von uns fordert. Die vielfältigen wissenschaftlichen Kontroversen und der langsame Fortschritt in den Wissenschaften belegen, dass dies keine evolutionären Anforderungen sind.

Fassen wir die obigen Überlegungen zusammen. Wissenschaftliches und kausales Denken zielen beide darauf ab, Kausalzusammenhänge zu erkennen. Unsere natürliche Umwelt und die klassische Naturwissenschaft stellen aber sehr unterschiedliche Anforderungen an unser Denken. Als biologische Organismen haben wir Jahrmillionen Zeit gehabt, uns an die Anforderungen unserer Umwelt anzupassen, als Wissenschaftler haben wir wenige Jahre Zeit, die Prinzipien wissenschaftlichen Arbeitens zu erlernen. Es verwundert daher nicht, wenn unser Denken den Anforderungen der Wissenschaft nur schlecht gerecht werden kann.

4.2.4 Denken mit und über Kausalmodelle - Reprise

Vom Standpunkt der Wissenschaft aus betrachtet ist sowohl das Denken mit als auch das Denken über Kausalmodelle entscheidend für einen Erkenntnisfortschritt. Beides wurde in dieser Arbeit untersucht. Die Experimente erforderten zum Teil die Anwendung von Kausalmodellen und zum Teil die Prüfung von Kausalmodellen. Beiden Anforderungen muss ein normativ denkender Wissenschaftler gerecht werden. Betrachten wir aber die Experimente und die mit ihnen verbundenen Anforderungen unter einem evolutionspsychologischen Blickwinkel, dann ergibt sich ein vollkommen anderes Bild.

Im ersten Experiment sollten die Probanden den Kausalzusammenhang zwischen zwei Ereignissen bestimmen. Zusätzlich lag ein drittes Ereignis vor, welches je nach Bedingung eine weitere Ursache oder ein weiterer Effekt der zu untersuchenden Ursache bzw. des zu untersuchenden Effektes war. Die Anforderung bei dieser Aufgabe war es also, einen Kausalzusammenhang über Beobachtungsdaten zu bestimmen und dabei eine mögliche Konfundierung mit einer zweiten Ursache auszuschalten. Diese Anforderung entspricht denen, welche auch unsere Umwelt an uns stellt. Die Aufgabe wich aber insofern von den natürlichen Gegebenheiten ab, dass die Daten in Form einer Liste ohne Informationen über die temporale Abfolge der Ereignisse vorgelegt wurden. Das Fehlen der zeitlichen Information war aus experimentaltechnischen Gründen notwendig. So konnte ein einheitlicher Datensatz in allen Bedingungen verwendet werden. Trotz dieser Abweichung bestimmten die Probanden den normativ richtigen und für den tatsächlichen Kausalzusammenhang indikativen statistischen Zusammenhang. Wenn das dritte Ereignis eine zweite Ursache des Effektes war, kontrollierten die Probanden den Einfluss dieses Ereignisses, so wie es für die korrekte Bestimmung des Kausalzusammenhangs notwendig ist.

In den Experimenten 2-5 wurde den Probanden die Aufgabe gestellt zu überprüfen, ob ein vermutetes Kausalmodell gegeben ist oder nicht. Dies ist eine typische Anforderung, welche die Wissenschaft an uns und unser Denken stellt. Es ist keine Anforderung, mit der wir in der Welt normalerweise konfrontiert werden. Die Daten wurden den Probanden einerseits als Listen, andererseits als Bilder dargeboten. In keinem Fall erhielten die Probanden Informationen über die Abfolge der Ereignisse. Dies war notwendig, um nicht über die zeitliche Abfolge Informationen über die Kausalstruktur zu vermitteln. Das wichtigste Ergebnis dieser Experimente war, dass die Probanden nicht richtig zwischen verschiedenen kausalen Strukturhypothesen differenzieren konnten.

Die Experimente 6-8 stellten innerhalb desselben Experimentes zwei verschiedene Anforderungen an die Probanden. Einerseits sollten sie ihr Wissen über das Modell und die Stärke der erlernten Kausalzusammenhänge nutzen, um Vorhersagen für Einzelfälle zu

machen. Diese Aufgabe entspricht weitgehend Anforderungen, wie wir sie aus unserem Alltag kennen. Die zweite Aufgabe war es, explizite Fragen über das Kausalmodell und seine strukturellen Implikationen zu beantworten. Dies ist eine wichtige Anforderung der Wissenschaft dafür, Modelle aus Daten abzuleiten und zu überprüfen. Das wichtigste Ergebnis dieser Studien war, dass die Probanden bei ihren Einzelvorhersagen sensitiv für die strukturellen Implikationen waren. Bei ihren expliziten Urteilen über den Zusammenhang der nicht direkt kausal verbundenen Ereignisse zeigten die Probanden keine Sensitivität für strukturelle Implikationen. Vermutlich haben Menschen kein deutliches, explizites Verständnis für die strukturellen Implikationen von Kausalmodellen. Dies ist auch keine evolutionäre Anforderung. Für das Lernen von Kausalzusammenhängen ist diese Art von Wissen nicht erforderlich, ebenso wenig wie für die Vorhersage von anderen Ereignissen.

Betrachten wir diese Befunde aus dem Blickwinkel evolutionärer Anforderungen, so kann man vereinfacht sagen, dass die Probanden immer dann eine sehr gute Leistung zeigten, wenn die Aufgabe evolutionären Anforderungen entsprach. Entsprachen die Anforderungen aber denen der Wissenschaft und verlangten ein Denken über Kausalmodelle, dann waren die Leistungen eher schlecht. Bemerkenswert ist, dass die Probanden versuchten die Strukturhypothesen durch die Bestimmung der Kausalzusammenhänge zu überprüfen. Sie setzten also ein Verfahren ein, das für eine andere, evolutionär wichtige Aufgabe angemessen und erfolgreich ist. Für die wissenschaftliche Aufgabe einer Strukturprüfung war das Vorgehen aber unangemessen.

Diese Interpretation geht natürlich weit über das hinaus, was aufgrund der vorliegenden Befunde als belegt gelten kann. Als Spekulation ist sie dennoch interessant, baut sie doch eine Brücke zwischen einer Psychologie kausalen Denkens, die auf normativen Modellen aufbaut, und einer, die auf evolutionspsychologischen Ansätzen basiert. Welche dieser Spekulationen durch empirische Forschung Unterstützung finden werden, muss die Zukunft zeigen.

LITERATUR

- Abramson & Finizza (1995). Probabilistic forecasts from probabilistic models. Case study: The oil market. *International Journal of Forecasting*, 11, 63-72.
- Ahn W., & Bailenson, J. (1996). Causal attribution as search for underlying mechanisms: An explanation of the conjunction fallacy and the discounting principle. *Cognitive Psychology*, 31, 82-123.
- Ahn, W., Kalish C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation vs. mechanism information in causal attribution. *Cognition*, 54, 299-352.
- Ahn, W., & Dennis, M. J. (2000). Induction of causal chains. In L. R. Gleitman and A. K. Joshi (Eds.). *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (S.19-24). Mahwah: Erlbaum.
- Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, 91, 112-149.
- Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, 108, 441-485.
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (1994). *Multivariate Analysemethoden*. Berlin: Springer.
- Baker, A. G., Mercier, P., Vallee-Tourangeau, F., Frank, R., & Pan, M. (1993). Selective associations and causality judgments: Presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 414-432.
- Birbaumer, N., & Schmidt, R. F. (1990). *Biologische Psychologie*. Berlin: Springer
- Bortz, J. (1990). *Statistik für Sozialwissenschaftler*. Berlin: Springer
- Buehner, M. J., & Cheng, P. W. (1997). Causal induction: The power PC theory versus the Rescorla-Wagner model. In the *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (S. 55-60). Hillsdale: Erlbaum.
- Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.). *Developmental psychology of time* (S. 209-254). New York: Academic Press.
- Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.

- Chapman, L. J. & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74*, 271-280.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367-405.
- Cheng, P. W., & Holyoak, K. J. (1986). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 392-416.
- Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality & Social Psychology, 58*, 545-567.
- Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99*, 365-382.
- Cheng, P. W., & Novick, L. R. (in Vorbereitung). Assessing interactive causal influence.
- Clemen, R. T. (1996). *Making hard decisions*. Pacific Grove: Duxbury Press.
- Davison, M. & McCarthy, D. (1988). *The matching law: A research review*. Hillsdale: Erlbaum.
- Dawes, R. (1988). *Rational choice in an uncertain world*. San Diego: Harcourt Brace Jovanovich.
- Dennis, M. J., & Ahn, W. (subm.). Primacy in causal strength judgments: The effect of initial evidence for generative vs. inhibitory relationships.
- Dörner, D. (1989). *Die Logik des Misslingens: Strategisches Denken in komplexen Situationen*. Hamburg: Rowohlt
- Domjan, M. (1998). *The principles of learning and behavior*. Pacific Grove: Brooks-Cole.
- Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17*, 397-434.
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.). *Judgment under uncertainty: Heuristics and biases* (S. 249-267). New York: Cambridge University Press.
- Edwards, W. (1998). Hailfinder: Tools for and experiences with Bayesian normative modelling. *American Psychologist, 53*, 416-428.
- Eells, E. (1991). *Probabilistic Causality*. Cambridge: Cambridge University Press.
- Evans, A. S. (1993). *Causation and disease: A chronological journey*. New York: Plenum.

- Fiedler, K., Walther, E., & Nickel, S. (2000). Covariation based attributions: On the ability to assess multiple covariates of an effect. *Personality and Social Psychology Bulletin*, 25, 607-622.
- Fiedler, K., Walther, E., & Nickel, S. (in Vorbereitung). Inductive social cognition: Experiments on the Simpson Paradox.
- Gallistel, R. & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, 107, 289-344.
- Gigerenzer, G., & the ABC Research Group (2000). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684-704.
- Gigerenzer, G. & Hug, K. (1992). Domain specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, 43, 127-171.
- Glymour, C. & Cooper, G. F. (1999). *Computation, causation and discovery*. Cambridge: MIT Press.
- Hempel, C. (1965). *Aspects of scientific explanations and other essays in the philosophy of science*. New York: Free Press.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24, 1-55.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge: MIT Press.
- Hume, D. (1987). *A treatise of human nature*. Oxford: Clarendon Press.
(Original aus dem Jahre 1739)
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Jensen, F. V. (1997). *An introduction to Bayesian networks*. London: Springer.
- Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs*, 7, 1-17.
- Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, 63, 305-400.
- Jungermann, H., Pfister, H.-R., & Fischer, K. (1998). *Psychologie der Entscheidung*. Heidelberg: Spektrum Akademischer Verlag.
- Kahneman, D., & Tversky, A (1982). The simulation heuristic. In D. Kahneman, P. Slovic and A. Tversky (Eds.). *Judgment under uncertainty: Heuristics and biases* (S. 201-208). New York: Cambridge University Press.

- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campell and R. M. Church (Eds.). *Punishment and aversive behaviour* (S. 279-296). New York: Appleton-Century-Crofts
- Kao, S. F., & Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 1363-1386.
- Kaiser, M.K., Profitt, D.R., & McCloskey, M. (1985). Development of beliefs about falling objects. *Perception and Psychophysics*, *38*, 533-539.
- Keinath, A. & Krems, J. F. (1999). Anomalous data integration in diagnostic reasoning. In A. Bagnara (Ed.). *Proceedings of the 1999 European Conference on Cognitive Science* (S. 213-218). Siena: Consiglio Nazionale delle Ricerche.
- Kelley, H.H. (1967). Attribution theory in social psychology. In D. Levine (Ed.). *Nebraska Symposium on Motivation*, *15* (S. 192-238). Lincoln: University of Nebraska Press.
- Kelley, H. H. (1972). Causal schemata in the attribution process. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins and B. Weiner (Eds.). *Attribution: Perceiving the cause of behavior*. Morristown: General Learning Press.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*, 107-128.
- Koslowski, B. (1996). *Theory and evidence. The development of scientific reasoning*. Cambridge: MIT Press.
- Klahr, D. (2000). *Exploring science. The cognition and development of discovery processes*. Cambridge: MIT Press.
- Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211-228.
- Krems, J. F., & Johnson, T. (1995). Integration of anomalous data in multi causal explanations. In J. D. Moore and J. F. Lehman (Eds.). *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (S. 277-282). Hillsdale: Erlbaum.
- Krist, H., Fieberg, E.F., & Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1-15.
- Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, *60*.
- Lien, Y., & Cheng, P. W. (2000). Distinguishing genuine from spurious causes: A coherence hypothesis. *Cognitive Psychology*, *40*, 87-137.

- Lober, K., & Shanks, D. R. (2000). Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, 107, 195-212.
- Loehlin, J. C. (1998). *Latent variable models. An introduction to factor, path, and structural analysis*. Mahwah: Erlbaum.
- Mackie, J.L. (1974). *The cement of the universe*. Oxford: Clarendon Press.
- McKenzie, C. R. M. (1994). The accuracy of intuitive judgment strategies: Covariation assessment and Bayesian inference. *Cognitive Psychology*, 26, 209-239.
- Michotte, A. E. (1963). *The perception of causality*. New York: Basic Books.
- Mill, J. S. (1968). *System der deduktiven und induktiven Logik*. Aalen: Scientia Verlag. (Original aus dem Jahre 1843, Übersetzung 1884)
- Oestermeier, U., & Hesse, F. W. (2000). Verbal and visual causal arguments. *Cognition*, 75, 65-104.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufman Publishers.
- Pearl, J. (1999). Probabilities of causation. Three counterfactual interpretations and their identification. *Technical Report*, University of California Los Angeles.
- Pearl, J. (2000). *Causality: models, reasoning, and inference*. Cambridge: Cambridge University Press.
- Plach, M. (1998). *Prozesse der Urteilsrevision*. Wiesbaden: Deutscher Universitäts Verlag
- Price, P. C., & Yates, J. F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition*, 21, 561-572.
- Reips, U.-D. (1997). *Erwerb von Kausalwissen: Lernrichtung und Diagnostizität von Effekten*. Unveröffentlichte Dissertation, Tübingen.
- Reips, U.-D., & Waldmann, M. R. (in Vorbereitung). Base rate appreciation after predictive and diagnostic learning.
- Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations on the effectiveness of reinforcement. In A. H. Black and W. F. Prokasy (Eds.). *Classical conditioning II: Current research and theory* (S. 64-99). New York: Appleton-Century-Crofts.
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, 133-148.
- Salmon, W. C. (1980). Probabilistic causality. *Pacific Philosophical Quarterly*, 61, 50-74.

- Schaller, M. (1992). In-group favoritism and statistical reasoning in social inference: Implications for formation and maintenance of group stereotypes. *Journal of Personality and Social Psychology*, 63, 61-74.
- Schaller, M. & O'Brian, M. (1992). 'Intuitive analysis of covariance' and group stereotype formation. *Personality and Social Psychology Bulletin*, 18, 776-785.
- Schauble, L. (1996). The development of scientific reasoning in knowledge rich contexts. *Developmental Psychology*, 32, 101-119.
- Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, 110, 101-120.
- Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Ser. B*, 13, 238-241.
- Shaklee, H., & Mims, M. (1982). Sources of error in judging event covariations: Effects of memory demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 208-224.
- Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory and Cognition*, 8, 459-467.
- Shanks, D. R. (1993). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology*, 84, 319-354.
- Shanks, D. R. , Lopez, F. J., Darby, R. J., & Dickinson, A. (1996). Distinguishing associative accounts and probabilistic contrast theories of of human contingency judgment. In D. R. Shanks, K. J. Holyoak and D. L. Medin (Eds.). *The Psychology of Learning and Motivation*, 34: *Causal learning* (S. 265-312). San Diego: Academic Press.
- Shanks, D. R. & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.). *The Psychology of Learning and Motivation*, 21 (S. 229-261). New York: Academic Press.
- Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, 47, 1-51.
- Spelke, E. S., Breilinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99, 605-632.
- Spellman, B. A. (1996a). Conditionalizing causality. In D. R. Shanks, K. J. Holyoak and D. L. Medin (Eds.). *The Psychology of Learning and Motivation*, 34: *Causal learning* (S. 167-207). San Diego: Academic Press.
- Spellman, B. A. (1996b). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science*, 7, 337-342.
- Spellman, B. A., & Mandel, D. R. (1999). When possibility informs reality: Counterfactual thinking as a cue to causality. *Current Directions in Psychological Science*, 8, 120-123.

- Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. New York: Springer.
- Thagard, P. (1999). *How scientists explain disease*. Princeton: Princeton University Press.
- Waldmann, M. R. (1994). *Der Erwerb von Kausalwissen: Prädiktives und diagnostisches Lernen mit Kausalmodellen*. Unveröffentlichte Habilitationsschrift, Tübingen.
- Waldmann, M. R. (1996). Knowledge based causal induction. In D. R. Shanks, K. J. Holyoak and D. L. Medin (Eds.). *The Psychology of Learning and Motivation, 34: Causal learning* (S. 47-88). San Diego: Academic Press.
- Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 53-76.
- Waldmann, M. R. (in press). Predictive versus diagnostic learning: Evidence from an overshadowing paradigm. *Psychological Bulletin & Review*.
- Waldmann, M. R., & Hagmayer, Y. (1995). When a cause simultaneously produces and prevents an effect. In J. D. Moore and J. F. Lehman (Eds.). *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (S. 425-430). Hillsdale: Erlbaum.
- Waldmann, M. R., & Hagmayer, Y. (subm.). Estimating causal strength. The role of structural knowledge and processing effort.
- 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.
- Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General, 124*, 181-206.
- Waldmann, M. R., & Martignon, L. (1998). A Bayesian network model of causal learning. In M. A. Gernsbacher and S. J. Derry, *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (S. 1102-1107). Mahwah: Erlbaum.
- Ward, W. C. & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology, 19*, 231-241.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129-140.
- Wason, P.C. (1968). On the failure to eliminate hypotheses in a conceptual task- A second look. In P. C. Wason and P. N. Johnson-Laird (Eds.). *Thinking and reasoning* (S. 165-174). Harmondsworth: Penguin.

-
- Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: The role of probability in judgments of response outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 174-188.
- Wasserman, E. A., Kao, S. F., Van Hamme, L. J., Katagiri, M., & Young, M. E. (1996). Causation and association. In D. R. Shanks, K. J. Holyoak and D. L. Medin (Eds.). *The Psychology of Learning and Motivation, 34: Causal learning* (S. 208-264). San Diego: Academic Press.
- Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, *56*, 161-169.
- White, P. A. (2000). Causal judgment from contingency information: Relation between subjective reports and individual tendencies in judgment. *Memory & Cognition*, *28*, 415-426.
- Wilkening, F., & Krist, H. (1995). Entwicklung der Wahrnehmung und Psychomotorik. In R. Oerter und L. Montada (Eds.). *Entwicklungspsychologie* (S. 487-517). Weinheim: Psychologie Verlags Union.
- Wu, M., & Cheng, P. W. (1999). Why causation need not follow from statistical association: Boundary conditions for the evaluation of generative and preventive causal powers. *Psychological Science*, *10*, 92-97.