

## 2. Kapitel

### Kausales Denken

*York Hagnauer und Michael R. Waldmann*

#### *1 Einführung: Konzepte von Kausalität*

Der Erwerb und die Nutzung von Kausalwissen gehört zu den grundlegenden kognitiven Kompetenzen, ohne die wir nicht überleben könnten. Kausale Zusammenhänge ermöglichen es uns, zukünftige Ereignisse vorherzusagen oder das Vorliegen von Ereignissen zu erklären. So lernen wir vorherzusagen, wie wir uns fühlen werden, wenn wir zu viel Alkohol getrunken haben, oder wir erklären uns unsere Magenverstimmung durch das Essen, das wir zuvor zu uns genommen haben. Kausalität betrifft nicht nur unsere eigenen Handlungen, sondern auch viele andere Bereiche unseres Alltags. Daher ist die Nutzung von Kausalwissen als eine bereichsbergreifende zentrale Kompetenz anzusehen. So können wir mithilfe von Kausalwissen Vorhersagen in Bereichen wie Physik (z. B. Flugbahn von Objekten), Biologie (z. B. Verhalten von Pflanzen und Tieren), Medizin (z. B. Krankheitsdiagnosen), Psychologie (Erklärung von Verhalten) und in vielen anderen Bereichen mehr machen. Kausalität wird in ganz unterschiedlichen Disziplinen erforscht. Viele wissenschaftliche Disziplinen versuchen, Gesetzmäßigkeiten in der Natur zu entdecken, die häufig auf kausalen Beziehungen beruhen. Demgegenüber interessiert sich die psychologische Kausalitätsforschung dafür, wie Menschen im Alltag Wissen über Kausalzusammenhänge erwerben und wie sie einmal erworbenes Wissen für Vorhersagen und Erklärungen nutzen (vgl. Spelber, Premack & Premack, 1995). Die beiden Aspekte Lernen und Denken sind in der Kausalitätsforschung eng verknüpft, weshalb wir sie auch zusammen behandeln werden.

### 1.1 Philosophische Theorien der Kausalität

Es mag verwundern, dass gleich am Anfang dieses Kapitels ein Abschnitt zu philosophischen Ansätzen steht. Dafür gibt es mehrere gute Gründe: Innerhalb der Philosophie, insbesondere in der Erkenntnistheorie, wurden zu Kausalität seit Jahrhunderten ähnliche Fragen gestellt, die auch heutige psychologische Ansätze stellen: *Wie werden kausale Zusammenhänge erkannt? Welche Annahmen sind dafür notwendig? Wie können kausale von nicht-kausalen Zusammenhängen unterschieden werden? Die Antworten auf diese Fragen bilden interessante und wohlurchdachte theoretische Ansätze, die sich auch in psychologische Hypothesen übersetzen lassen. Viele psychologische Theorien haben Vorläufer in der Philosophie. Dies soll natürlich nicht heißen, dass innerhalb der Philosophie keine eigenständigen Fragen gestellt werden. Das Gegenteil ist der Fall. Aktuelle Beispiele hierfür sind Fragen nach der Beziehung von freiem Willen und kausalem Determinismus oder die kausale Interpretation quantenphysikalischer Ansätze in der Physik. In diesem Kapitel werden wir uns aber auf philosophische Ansätze beschränken, welche die Theoriebildung in der Psychologie betrücht haben. Trotz dieser begrenzten Auswahl können die jeweiligen Theorien nur verkürzt, konzentriert auf wenige Aspekte dargestellt werden.*

Die Frage nach Kausalität, das heißt nach den Ursachen des Geschehens in der Welt, beschäftigt die Philosophie seit Anbeginn. Bereits die beiden Vorsokratiker Heraklit und Demokrit haben das Kausalitätsprinzip formuliert, welches besagt, dass nichts in der Welt ohne Ursache geschieht („*nihil fit sine causa*“, vgl. Audi, 1995). Die Jahrtausende überdauernde Akzeptanz dieses Prinzips spricht auch dafür, dass es psychologisch sehr plausibel ist.

Aristoteles stellte die These auf, dass es vier Arten von Ursachen gibt, die stofflich-materielle Ursache („*causa materialis*“), die Form-Ursache („*causa formalis*“), die Wirkursache („*causa efficiens*“) und die Ziel-Ursache („*causa finalis*“). In der Folge wurde der Ursachebegriff langsam auf die Wirkursache eingeschränkt. Seit Beginn der Neuzeit werden unter Ursache Ereignisse oder Zustände verstanden, die ein anderes Ereignis hervorrufen.

Mehr als andere Philosophen hat der schottische Philosoph David Hume das moderne Verständnis von Kausalität geprägt. In Humes Werk liegt die in der aktuellen Kognitionspsychologie dominierte Verknüpfung von Assoziations-theorien mit der Theorie der Kausalität begründet. Er mag damit als Urvater heutiger Versuche gelten, den Erwerb von Kausalwissen auf psychologische Prinzipien assoziativen Denkens zu reduzieren.

Aufbauend auf die Überlegungen von Aristoteles in seinem Werk „*Para Naturalia*“ nennt Hume (1739/1978, 1748/1977) drei Prinzipien, auf die sich seiner

Meinung nach alle Assoziationen zurückführen lassen: 1. Ähnlichkeit („*resemblance*“), 2. raum-zeitliche Kontiguität („*contiguity*“), 3. Kausalität („*cause and effect*“). Er unterscheidet also zwischen Assoziationen, die auf raum-zeitlicher Kontiguität beruhen und deshalb auch bei *Einzelfällen* gebildet werden können, und solchen, die auf einem Ursache-Wirkungseindruck beruhen und damit seiner Theorie zufolge Mehrfachbeobachtungen voraussetzen. Während die Logik die Grundzüge deduktiven Denkens („*reason*“) bildet, sei Kausalität die Grundraum-zeitlich beobachtbar auftreten können, geht ein Kausalendruck über die singuläre Wahrnehmung von zwei Ereignissen hinaus. Das gemeinsame Auftreten von Ursache und Effekt wird nämlich von dem Eindruck der Notwendigkeit begleitet. Eine der zentralen Fragen von Humes Philosophie ist nun, woher dieser Eindruck der Notwendigkeit rührt.

Die traditionelle Antwort wäre damals gewesen, dass Ursachen Ereignisse sind, denen eine kausale Potenz („*causal power*“) innewohnt, die sie auf Effekte übertragen. Hume stellt die Frage, wie wir die kausale Potenz der Ursache bzw. die Notwendigkeit der kausalen Verbindung von zwei Ereignissen („*necessary connection*“) erkennen. Wie alle britischen Empiristen geht er davon aus, dass Wissen über die Welt allein auf der Erfahrung („*experience*“) singulärer Ereignisse beruht und dass es keine angeborenen Ideen („*innate ideas*“) gibt, wie sie etwa von Descartes (1972) postuliert wurden. In der Sprache der modernen Kognitionspsychologie würde man sagen, dass Humes und Lockes Auffassung zufolge Wissen rein datergesteuert („*bottom up*“) erworben wird. Der Organismus sei eine *tabula rasa* (Locke, 1700), bevor er zum ersten Mal mit Erfahrungen konfrontiert wird.

Humes Empirismus zusammen mit seiner epistemischen Grundannahme, dass unsere Erfahrungen sich auf *Ereignisse* beziehen, führt ihn zu der Schlussfolgerung, dass alles, was wir beim Eindruck von Kausalität wahrnehmen können, raum-zeitlich beobachtbare Ereignisse sind. Es gibt nichts in unseren Wahrnehmungen, was direkt der Notwendigkeit kausaler Beziehungen entspricht.

Wie entsteht nun der Eindruck kausaler Power? Humes Antwort ist, dass es sich um ein Produkt psychologischer Mechanismen handelt. Nimmt man wieder die Sprache der aktuellen Kognitionspsychologie zu Hilfe, dann kann man sagen, Humes These war, dass es sich um eine *kognitive Illusion* handelt. Ein Kausalendruck entsteht nicht beim einmaligen Erlebnis einer raum-zeitlichen Abfolge von Ereignissen, sondern erst nach wiederholter Erfahrung des gemeinsamen Auftretens ähnlicher Ereignispaare („*constant conjunction*“). Je häufiger die Paarung gesehen wird und je ähnlicher sich die Ereignisse sind, desto stärker ist der Kausalitätseindruck. Die Erfahrung von Kausalität ist damit nicht

das Produkt wahrgenommener kausaler Notwendigkeiten, sondern ein Ergebnis der Assoziationsstärke („habit“, „custom“). Hier sieht man, wie Humes Kausalität auf die bereits von Aristoteles postulierten Assoziationsprinzipien von raum-zeitlicher Kontiguität und Ähnlichkeit zurückführt. Als weitere Bedingung führt er an, dass Ursachen vor ihren Effekten auftreten, eine Annahme, die er nicht weiter begründet und die nicht so recht zu den anderen assoziationsorientierten Prinzipien passen mag. Hume (1739/1978) selbst fasst die drei wichtigsten Prinzipien von Kausalentdeckungen folgendermaßen zusammen:

1. The cause and effect must be contiguous in space and time.
2. The cause must be prior to the effect.
3. There must be a constant union betwixt the cause and effect. 'Tis chiefly by this quality, that constitutes the relation.“ (S. 173)

Insbesondere in seinem späteren Werk (1748/1977) leugnet Hume nicht, dass es so etwas wie notwendige Kausalbeziehungen in der Welt geben mag. Seine These ist nur, dass wir nicht in der Lage sind, Wissen über solche kausalen Prozesse zu erlangen. Alles was wir zur Verfügung hatten, sei unsere Fähigkeit Ereigniswiederholungen wahrzunehmen.

Humes Ideen wurden in der psychologischen Theoriebildung und Forschung vielfach aufgegriffen. Die Rolle der Kontiguität für die Bildung von Kausalurteilen wurde sowohl in entwicklungspsychologischen Arbeiten (z. B. Bullock, Gelman & Baillargeon, 1982; Schultz, 1982) als auch in Studien zur phänomenalen Kausalität untersucht (Michotte, 1946). Die Regularitätsannahme findet sich sowohl in assoziativen wie auch regelbasierten Ansätzen wieder.

Kant (1781) greift in seiner Kritik der reinen Vernunft die Ideen Humes auf und stimmt ihm darin zu, dass ein allgemein gültiges, das heißt notwendiges Kausalgesetz nicht empirisch begründbar ist. Die Annahme, dass Kausalzusammenhänge uniform in Vergangenheit wie Zukunft gleich wirksam sind, muss vielmehr aus dem Denken selbst stammen. Ja mehr noch, diese Annahme bildet nach Kant eine notwendige Voraussetzung dafür, dass überhaupt Ursache-Wirkungszusammenhänge in der Welt erkannt werden können. Wenn wir den Begriff der Kausalität nicht hätten, bevor wir Erfahrungen machen, so würden wir nur eine Vielzahl unzusammenhängender Ereignisse sehen. Der Begriff der Kausalität und andere Konzepte ordnen das beobachtete Geschehen erst für uns, wobei die Kausalität uns immer währende Sequenzen von Ereignissen liefert.

Kants Überlegungen wurden in der Psychologie ebenfalls aufgegriffen, wenn auch weit weniger als die Ideen Humes. Ein Grund hierfür mag seine idealistische empirierene Konzeption sein, ein anderer, dass seine Begrifflichkeit erst in psychologische Termini und Theorien übersetzt werden muss. Mir seiner Idee,

dass der Begriff der Kausalität notwendig für das Erkennen von Kausalität ist, wird Kant zum Pater der Theorien, die annehmen, dass Vorwissen über Kausalität die Wahrnehmung und die Beurteilung von Zusammenhängen als kausal beeinflusst. Hier ist insbesondere die Theorie der Kausalmodelle zu nennen, welche die Rolle abstrakten Kausalwissens betont (vgl. Waldmann, 1996). Auch Cheng (1997) beruft sich auf Kant, wenn sie die Annahme begründet, dass wir mit der Vorannahme an kausale Situationen herangehen, dass es Ursachen mit einer innen innewohnenden kausalen Potenz („power“) gibt.

Neuere philosophische Theorien entwickelten vor allem die Ideen Humes in unterschiedlichen Richtungen weiter. Ausgangspunkt für die verschiedenen Theorien ist der Zusammenhang von Ursache und Wirkung. Der ursprünglichen Konzeption von Hume zufolge sind Ursachen notwendig und hinreichend für den Effekt. Der Effekt tritt immer dann und nur dann auf, wenn die Ursache vorliegt. Empirische Beobachtungen zeigen aber, dass dies nicht immer zutrifft. Es muss eine bestimmte Menge von Bedingungen gegeben sein, damit ein Effekt eintritt. Der Effekt kann aber auch unter anderen Bedingungen eintreten. Beispielsweise kann ein Waldbrand dadurch entstehen, dass ein Blitz in den vertrockneten Wald einschlägt, aber auch, weil mittels eines Brandbeschleunigers ein Feuer gelegt wurde. Zwischen einem einzelnen Ursachefaktor und dem Effekt besteht also weder ein hinreichender noch ein notwendiger Zusammenhang. Dennoch gehen wir davon aus, dass es sich um Ursachen handelt. Warum? Neuere Kausalitätstheorien versuchen u. a. eine Antwort auf diese Frage zu geben. Eine zweite Frage, die sich im Anschluss an die Theorien von Hume und Kant stellt, ist, wodurch perfekt deterministische Zusammenhänge, die nicht kausal sind, von solchen unterschieden werden können, die kausal sind. Ein häufig diskutiertes Beispiel sind Tag und Nacht. Obwohl sie einander deterministisch folgen und alle Kriterien Humes erfüllt sind, wollen wir nicht behaupten, dass der Tag die Nacht hervorruft. Auch für dieses Problem liefern unterschiedliche Ansätze verschiedene Antworten.

In seinem bedingungsanalytischen Ansatz greift Mackie (1974) die obigen Analysen direkt auf und definiert Ursachen neu als so genannte INUS-Bedingungen („Insufficient but Necessary parts of Unnecessary but Sufficient conditions“), das heißt, Ursachen sind selbst nicht für den Effekt hinreichend, aber sie sind ein notwendiger Bestandteil einer Menge von Bedingungen, die in ihrer Gänze hinreichend für den Effekt sind. Dieser Ansatz vermag auch das zweite der oben geschilderten Probleme zu lösen. Da in einer vollständigen Bedingungsanalyse von Tag und Nacht die Sonne und die Erddrehung miteinbezogen werden müssen, ergibt sich, dass weder der Tag für die Nacht noch die Nacht für den Tag eine INUS-Bedingung ist. Mackie stelle auch eine Reihe von Spekulationen darüber an, welche der Bedingungen psychologisch gesehen als „die Ursache“ eines Ereignisses gekennzeichnet werden. So vermutete er unter anderem, dass eher un-

gewöhnliche oder willkürlich durch Handlungen hervorgerufene Ereignisse als Ursachen von Ereignissen herangezogen werden als regelmäßig vorkommende Ereignisse oder konstant vorliegende Randbedingungen. Mackies Analysen und psychologische Hypothesen wurden von Einhorn und Hogarth (1986) in ihrer Theorie der kausalen Urteilbildung aufgegriffen (vgl. Abschnitt 2.5).

Der kontrafaktische Ansatz von David Lewis (1973, 1986) greift eine zweite Ursachendefinition von Hume auf: „we may define a cause to be an object, followed by another, (...) or in other words, where if the first object had not been the second never had existed“ (Hume, 1739/1978, S. 76). Das bedeutet, ein Ereignis ist dann die Ursache eines Effektes, wenn der Effekt bei Ausbleiben der Ursache nicht eingetreten wäre. Diese Ursachendefinition verlangt nicht mehr, dass eine Ursache hinreichend für den Effekt ist, aber dass sie unter den gegebenen Umständen notwendig ist. Unter anderen Bedingungen ist dasselbe Ereignis möglicherweise nicht mehr notwendig für den Effekt. In diesem Fall ist es dann allerdings auch keine Ursache. Bei einer Brandstiftung ist der Brandbeschleuniger notwendig für das Entstehen eines Waldbrands und damit eine Ursache; innerhalb des Blitzschlag szenarios ist er es aber nicht. Damit löst diese Analyse elegant das erste der beiden Probleme: Echte von scheinbaren Ursachen können ebenfalls unterschieden werden. Scheinbare Ursachen sind im Sinne der Definition für den Effekt nicht notwendig. Auch wenn es nicht Nacht gewesen wäre (z. B. wegen der Mittsomernacht in der Nähe des Polarkreises), würde der Tag beginnen. In der Psychologie werden kontrafaktische Aussagen schon länger untersucht (z. B. Wells & Gavanski, 1989). Erst vor Kurzem wurde ihre Bedeutung in Bezug auf Kausalität näher beleuchtet (Sloman & Lagnado, 2002, 2005).

Handlungstheoretische Ansätze (z. B. v. Wright, 1971) sehen Kausalwissen primär als Handlungswissen an. Ursachen sind Handlungen, welche zweckorientiert sind und ein bestimmtes Ereignis, den inneren Effekt, hervorruft. Das Wissen über die Wirkung verschiedener Handlungen wird durch aktive Praxis erworben. Alle anderen Arten von Ursachen, die wir kennen, sind von den Handlungsursachen abgeleitet. So können beispielsweise die Sonne und die Erdrotation als Ursachen für Tag und Nacht auch über Handlungen verstanden werden. Wird eine Lichtquelle an ein rotierendes Objekt gehalten, dann beleuchtet das Licht immer nur die zu einem bestimmten Zeitpunkt der Lampe zugewandte Seite. Mittels des handlungstheoretischen Ansatzes lässt sich auch erklären, warum Ursachen weder notwendig noch hinreichend für einen Effekt sind. Meistens sind verschiedene Handlungen möglich, um einen gewissen Effekt hervorzurufen. Die Unterscheidung von echten im Gegensatz zu scheinbaren Ursachen ergibt sich ebenfalls direkt über das Konzept der Handlung. Kann durch eine Handlung ein Effekt beeinflusst werden, so handelt es sich um eine Ursache des Effektes. Führt die Manipulation eines Ereignisses nicht dazu, dass sich ein anderes Ereignis verändert, dann liegt kein Kausalzusammenhang

vor. Etwas beobachtete Kovariationen dieser Ereignisse sind also nicht durch eine direkte kausale Verbindung bedingt. Der handlungstheoretische Ansatz wurde manchmal kritisiert, weil er voraussetzen scheint, dass wir immer Ursachen durch Handlungen setzen können müssen. Dies ist aber natürlich nicht immer der Fall (z. B. bei Planetenbewegungen). Woodward (2003) hat deshalb den handlungstheoretischen Ansatz erweitert. Er analysiert nun Kausalität im Rückgriff auf kontrafaktische Annahmen über hypothetische Interventionen, die bestimmte kausale Eigenschaften haben müssen. Die reale Durchführbarkeit der Handlung ist in dieser Konzeption ein methodisches Plus, ist aber keine notwendige Voraussetzung für das konzeptuelle Verständnis von Kausalität.

Obwohl der handlungstheoretische Ansatz der Kausalität psychologisch sehr interessant und fruchtbar zu sein scheint, wurde er bis dato wenig beachtet. Dies verwundert umso mehr, als dass Forschungen sowohl zum wissenschaftlichen Denken als auch zum instrumentellen Konditionieren darauf hinweisen, dass die aktive Manipulation einer Ursache zu akkuraten Kausalurteilen führt. Dies ist aber bei einer reinen Beobachtung der Kovariation zweier Ereignisse nicht unbedingt der Fall.

Probabilistische Ansätze (Eells, 1991; Laplace, 1814/1912; Pearl, 1988; Salmon, 1984; Spirtes, Glymour & Scheines, 2001; Suppes, 1970, 1984) nehmen die Erkenntnis, dass zwischen Ereignissen in der Welt nur probabilistische Zusammenhänge vorliegen, zum Ausgangspunkt ihrer Überlegungen. Unterschiedliche Meinungen gibt es bei Vertretern dieser Ansätze darüber, ob die Kausalzusammenhänge *per se* probabilistisch oder aber eigentlich deterministisch sind und nur in unserer Erfahrung probabilistisch erscheinen. Ein Vertreter letzterer Auffassung war Laplace (1814/1912) und in jüngster Vergangenheit Pearl (2000). Diese behaupten, dass der Eindruck, dass Kausalzusammenhänge probabilistisch sind, deshalb zu Stande kommt, weil zu einem bestimmten Zeitpunkt stets eine Fülle gleichzeitig wirkender Ursachen, von denen die meisten auch noch unbekannt sind, ihre Wirkung entfalten. Trotz dieser Unterschiede in der metaphysischen Konzeption von Ursachen stimmen Vertreter beider Richtungen darin überein, dass Ursachen auf Grund der probabilistischen Zusammenhänge erkannt werden können. Ursachen zeichnen sich darin aus, dass bei ihrem Vorliegen die Wahrscheinlichkeit des Eintretens des Effektes erhöht wird (generative Ursachen) oder aber verringert wird (inhibitorische Ursachen). Verglichen wird dabei jeweils die Wahrscheinlichkeit des Effektes bei Vorliegen der Ursache ( $P(E|U)$  Ursache) =  $P(E|\bar{U})$  mit der Wahrscheinlichkeit des Effektes ohne Vorliegen der Ursache ( $P(E)$ Keine Ursache) =  $P(E|\bar{c})$ ). Die Differenz der beiden Wahrscheinlichkeiten wird dabei auch als Kontingenz oder  $\Delta P$  bezeichnet. In der Sprache der Wahrscheinlichkeitstheorie wird diese folgendermaßen definiert:

$$\Delta P = P(E|c) - P(E|\bar{c}) \text{ (Kontingenzformel)}$$

Generative Ursachen haben eine positive Kontingenz, inhibitorische Ursachen eine negative.

Auch probabilistische Ansätze sind in der Lage, echte von scheinbaren Kausalzusammenhängen zu unterscheiden. Scheinbar wie echte Kausalzusammenhänge zweiter Ereignisse weisen eine positive Kontingenz auf. Bei scheinbaren Ursachen beruht diese Kontingenz aber nicht auf einem Kausalfluss dieser Ursache, sondern auf der Wirkung einer anderen Ursache, welche beide Ereignisse, die scheinbare Ursache und den Effekt, beeinflusst. In unserem Tage- und Nachtbeispiel ist diese gemeinsame Ursache die Erdrotation in Verbindung mit der Sonneneinstrahlung. Wird der kausale Einfluss dieser Ursachen kontrolliert, so verschwindet der scheinbare Kausalzusammenhang. Ohne Erdrotation oder ohne Sonneneinstrahlung gäbe es keine Abfolge mehr von Tag und Nacht. Aufbauend auf diese Überlegung, lässt sich eine Definition von echten Kausalursachen angeben. Ein Ereignis ist eine echte Kausalursache, wenn die Wahrscheinlichkeit des Effektes bei Anwesenheit des Ursacheereignisses im Vergleich zur Abwesenheit der Ursache auch dann noch erhöht ist, wenn alle anderen Ursachen kontrolliert werden (vgl. Cartwright, 1983; Eells, 1991). Die Theorie der probabilistischen Kausalität hat viele Untersuchungen in der Psychologie stimuliert, die probabilistische Kontingenzen als Ausdruck von Kausalität gesehen haben (z. B. Cheng & Novick, 1990, 1992).

Trotz vielversprechender Anfänge gilt der Versuch, Kausalität allein auf probabilistische Relationen zwischen Ereignissen zu reduzieren, als weitgehend gescheitert (vgl. Pearl, 2000; Salmon, 1984). Ein Problem besteht darin, dass kausale Beziehungen gerichtet sind (von Ursachen zu Effekten), sich die Richtung aber nur in bestimmten Fällen und unter Zuhilfenahme von Zusatzannahmen aus probabilistischen Daten rekonstruieren lässt. Ein weiteres Problem sind die potenziell konfundierenden Ko-Faktoren. Eine statistische Korrelation deutet nur dann auf einen Kausalzusammenhang, wenn *alle* konfundierenden Faktoren konstant gehalten werden. Woher soll man aber wissen, dass man alle Faktoren kennt und woher soll man wissen, dass es sich um relevante alternative Kausalfaktoren handelt und nicht um irrelevante Ereignisse (vgl. Cartwright, 1989)? Dies erzeugt aber ein Zirkularitätsproblem. Am Ende muss man bereits über Kausalwissen verfügen, um neues Kausalwissen zu erhalten. Waldmann und Hagemyer (2001) haben in Experimenten gezeigt, dass Versuchsteilnehmer tatsächlich nur Faktoren als Ko-Faktoren berücksichtigen, die sie vorab für kausal relevant halten.

Manche Frühen Anhänger des Programms der probabilistischen Kausalität haben deshalb dieses Projekt fallen gelassen und haben versucht, die jahrhundertalte Konzeption von Kausalität als Übertragung von Energie, einer physikalischen Quantität oder Information zu modernisieren (vgl. Dowe, 2000; Salmon, 1984).

Kausale Beziehungen sind beispielsweise dem Ansatz von Salmon zufolge Prozesse, bei denen eine physikalische Größe („mark transmission“) übertragen wird. Davon zu unterscheiden sind Pseudo-Prozesse. Während eine Lichtquelle einen kausalen Prozess initiiert, den man auf dem Weg unterbrechen kann (etwa indem man den Lichtstrahl mit einem Blatt Papier blockiert), lassen sich Pseudo-Prozesse (bspw. der bewegende Lichtpunkt an der Wand) nicht auf diese Weise unterbrechen. Pseudo-Prozesse stellen also keine kausalen Prozesse dar. Dieser Ansatz wurde ausgearbeitet, um auch komplexere Kausalstrukturen mit mehreren Effekten und interagierenden Ursachen zu beschreiben. In der Psychologie, insbesondere der Entwicklungspsychologie, wurde die Konzeption, dass es bei kausalen Beziehungen um Mechanismen geht, die Energie übertragen, sehr populär (vgl. Ahn, Kalish, Medin & Gelman, 1995; Bullock, Gelman & Baillargeon, 1982; Shultz & Kestenbaum, 1985).

Nur wenige Philosophen haben sich diesem Programm angeschlossen. Zu sehr scheint es dem Bereich physikalischer Kausalität verhaftet. Außerdem lässt sich diese Konzeption nur anwenden, wenn man bereits über Detailwissen über kausale Prozesse verfügt. Aber gerade die Fragen, wie wir neues Kausalwissen auf der Basis von Beobachtungen erwerben und wie wir Kausalwissen repräsentieren, über dessen Mechanismen wir nichts wissen, lassen sich mit diesem Ansatz nicht beantworten.

Neuere probabilistische Kausalitätstheorien (Pearl, 2000; Glymour & Cooper, 1999; Spirtes et al., 2001; Woodward, 2003) haben sich mit der Frage beschäftigt, wie und unter welchen Bedingungen sich aus vorliegenden probabilistischen Daten über die Zusammenhänge von Ereignissen die Struktur der zu Grunde liegenden Kausalzusammenhänge rekonstruieren lässt und haben hierfür auch Computermodelle entwickelt („kausale Bayes-Netze“). Dieser Ansatz greift die Tradition der probabilistischen Kausalität auf; geht aber davon aus, dass abstraktes Vorwissen über Kausalmodelle, wahrnehmbare Hinweise („Cues“) (z. B. die temporale Reihenfolge der Ereignisse), probabilistische Informationen und einschänkernde Zusatzannahmen zusammenspielen müssen, um zu adäquatem Kausalwissen zu gelangen. Damit geht es nicht um Induktion auf der Basis von probabilistischer Information allein, sondern eher um Konsistenzprüfung zwischen kausalen Modellen und den Daten. Daten können zwar zwischen Modellen entscheiden und auch die Bildung von neuen Theorien stimulieren, aber nur, wenn man bestimmte Vorannahmen macht.

Die Theorie der kausalen Bayesnetze und verwandte Ansätze wurden von psychologischen Kausalitätstheorien stark rezipiert (Cheng, 1997; Glymour, 2001, 2003; Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Gopnik, Sobel, Schulz & Glymour, 2001; Waldmann, 1996; Waldmann & Hagemyer, 2001, 2005; Waldmann & Holyoak, 1992). Trotz ihrer Komplexität und formalen

Sterge haben sie sich als Quelle interessanterer neuartiger Hypothesen als extrem fruchtbar erwiesen. Sie bilden die Grundlage sowohl für die im Folgenden vorgestellten regelbasierten als auch für die kausalmodelltheoretischen Ansätze kausalen Denkens und Lernens.

Zusammenfassend ist zu sagen, dass heutige psychologische Theorien der Kausalität ohne die in der Philosophie geleisteten Vorarbeiten kaum denkbar wären. Im Bereich der probabilistischen Theorien der Kausalität zeichnet sich zudem eine enge Zusammenarbeit beider Disziplinen ab.

## 2 Psychologische Theorien kausalen Denkens und Lernens

Der Überblick über philosophische Ansätze zur Kausalität hat bereits die Breite der theoretischen Ansätze in der Psychologie deutlich gemacht. In diesem Kapitel wird eine Übersicht über die wichtigsten derzeit diskutierten Ansätze gegeben. Wir beginnen mit Assoziationsstheorien, die von der Grundannahme ausgehen, dass der Erwerb von Kausalwissen im Grunde das Gleiche ist wie klassisches und instrumentelles Konditionieren. Dieser Sicht zufolge werden keine statistischen Analysen vorgenommen, sondern es werden Reize, die beispielsweise Ursachen repräsentieren können, mit anderen Reizen assoziiert, die die Effekte beschreiben. Gelemt werden Assoziationsgewichte, die proportional zur Häufigkeit der Paarung dieser Reize und in Abhängigkeit weiterer Faktoren steigen oder sinken. Ein damit konkurrierender Ansatz sind regelbasierte Theorien, die davon ausgehen, dass wir im Alltag intuitive Satzstriker sind, die die Beobachtungen mithilfe von statistischen Regeln bindeln und diese Regeln dazu nutzen, auf kausale Zusammenhänge unterschiedlicher Stärke zu schließen. Eine Vielzahl von Theorien konkurriert hier darüber, welche Regeln wir verwenden, wenn wir aus Beobachtungen Wissen über Kausalzusammenhänge erwerben. Der deduktive Ansatz, der im dritten Abschnitt erläutert wird, liegt weitgehend quer zu den anderen Ansätzen. Hier geht es nicht um den Erwerb von Kausalwissen, sondern um das Schließen aus bereits vorgegebenen kausalen Prämissen. Es zeigt sich, dass sich diese Art des Schließens nicht mit traditionellen Theorien logischen Schließens erklären lässt. Im Anschluss an diesen Ansatz diskutieren wir eine Klasse von Theorien, die sich gegen die Sicht stellen, dass es bei Kausalität primär um die statistische Beziehung von Ereignissen oder die Assoziation zwischen Ereignissen geht. Die Mechanismus-Sicht stellt vielmehr die kausalen Mechanismen in den Vordergrund, die zwischen Ursache und Effekt vermitteln. Danach wenden wir uns einem urteilsheuristischen Ansatz zu, welcher davon ausgeht, dass bei Kausalurteilen stets mehrere verschiedene Indikatoren von Kausalität einbezogen und miteinander verrechnet werden. Schließlich werden wir neuere Ansätze ansprechen, die sich mit komplexeren Kausalmodellen befassen. Diese Theorien, die sich mithilfe von Bayes-Modell-

len darstellen lassen, stellen ein interessantes Bindeglied zwischen statistischen und mechanistischen Ansätzen dar.

Im Mittelpunkt der Darstellung werden Modelle und experimentelle Befunde stehen, die von Anhängern der jeweiligen Theorie zur Stützung ihres Ansatzes erbracht wurden. Im Anschluss daran folgt jeweils eine Diskussion von Kritikpunkten an den Theorien.

### 2.1 Assoziationsstheorien

Untersuchungen zum Erwerb von Kausalwissen verwenden häufig Aufgaben, in denen die Versuchsteilnehmer nacheinander Informationen über Anwesenheit oder Abwesenheit von Ursachen und Effekten erhalten („trial-by-trial learning“). Eine Lernaufgabe mag darin bestehen, dass Versuchsteilnehmer zunächst Informationen über Nahrungsmittel bekommen, die einzelne Personen zu sich genommen haben, und dann vorhersagen lernen, welche dieser Personen eine Allergie bekommen.

Prozedural sind solche Aufgaben sehr ähnlich den Versuchsanordnungen, die man bei Untersuchungen zum klassischen und instrumentellen Konditionieren verwendet (vgl. Domjan, 2003; Lachnit, 1993), so dass es nahe liegt, diese verschiedenen inhaltlichen Aufgaben als Manifestationen eines allgemeinen, reichsübergreifenden assoziativen Lernmechanismus anzusehen (vgl. Waldmann, 1996, 1996a, 1997). Genauso wie der Pawlow'sche Hund eben lernt, Tonreize mit Futter in Verbindung zu bringen, stellt man sich vor, dass die Versuchsteilnehmer in der eben erwähnten Anordnung lernen, Nahrungsmittel mit Allergien zu assoziieren.

Wir haben im ersten Abschnitt bereits gesehen, dass die Vorstellung, dass der Eindruck von Kausalität im Grunde eine Illusion ist, die genährt wird durch die Stärke von Assoziationen, auf Hume (1739/1978) zurückgeht. Hume ging davon aus, dass raum-zeitlich benachbarte Ereignisse („Kontiguitätsprinzip“), die regelmäßig zusammen auftreten, miteinander assoziiert werden. Die Stärke dieser Assoziation ist die Basis für unseren Eindruck der Stärke der Kausalbeziehung. Das Kontiguitätsprinzip hat lange Zeit auch psychologische Theorien des assoziativen Lernen dominiert. Erst seit etwa 35 Jahren mehren sich Befunde, die zeigen, dass Kontiguität allein nicht hinreichend ist, um assoziatives Lernen zu erklären. So hat Rescorla (1968) gezeigt, dass neben Kontiguität auch Kontingenzen entscheidend ist. Übertragen auf den Bereich Kausalität bedeutet dies, dass nicht nur die Anzahl der Paarungen von Ursache und Effekt entscheidend für die Assoziationsstärke ist, sondern auch die Anzahl der Ereignisse, in denen Effekte in Abwesenheit der Ursachen auftreten. Wenn wir etwa beobachten,

dass eine Allergie bei einer Person in 80% der Fälle auftritt, in denen sie Nüsse gegessen hat, aber dass dieselbe Person auch in 80% der Fälle eine allergische Reaktion hat, in denen sie keine Nüsse gegessen hat, dann würden wir Flüsse nicht verantwortlich machen für die Allergie (trotz der häufigen Paarung).

Ein weiteres Problem für traditionelle, konfigurationsbasierte Theorien des assoziativen Lernens sind Befunde, die zeigen, dass Hinweisreize im Wettbewerb miteinander sehen, so dass selbst häufig räumlich-zeitlich gepaarte Ereignisse nicht assoziiert werden („cue competition“). Der klassische empirische Befund, der kompetitives Lernen dieser Art belegt, ist der Blockierungseffekt (Kamin, 1968, 1969). Das zugehörige Blockierungsparadigma besteht aus zwei Lernphasen. In Phase 1 lernen die Tiere eine perfekte Beziehung (Kontingenz von 1) zwischen einem konditionierten Stimulus (CS1) und einem unkonditionierten Stimulus (US). In Phase 2 wird ein zweiter CS (CS2) mit dem ersten CS1 gepaart. Nun folgt auf beide CS der US. Die Ergebnisse zeigen, dass die Tiere nur die Beziehung zwischen dem CS1 und dem US gelernt zu haben scheinen, während der Erwerb von Wissen über den CS2 offenbar blockiert wurde. Dieser Blockierungseffekt tritt ein, obwohl zwischen dem CS2 und dem US ebenfalls eine perfekte Kontingenz von 1 vorliegt. Intuitiv erklären kann man diesen Befund damit, dass die Tiere nur dann eine assoziative Beziehung lernen, wenn sie dadurch andere Reize (den US in diesem Fall) besser vorhersagen können. Da der CS2 aus Phase 2 die Vorhersage des US relativ zu dem bereits gelernten perfekten Prädiktor CS1 nicht verbessert, findet kein Lernen statt.

Dieser und andere ähnliche Befunde haben zu einer fundamentalen Veränderung der Theorien assoziativen Lernens geführt. Diese neuen Theorien weisen über das Lernen von kausalen Kontingenzen zwischen Ereignissen hinaus herkömmlichen Vorstellungen mechanischen, konfigurationsbasierten Lernens (vgl. Rescorla, 1991). Es kann deshalb nicht erstaunen, dass man erst recht die Normendigkeit einer eigenständigen Theorie des kausalen Lernens in Frage stellt wurde. So haben Dickinson, Shanks und Evident (1984) als Erste postuliert, dass kausales Lernen und assoziatives Lernen, so wie es die modernen Theorien modellieren, sich nicht nur oberflächlich ähnlich sind, sondern dass beide durch einen gemeinsamen assoziativen Lernmechanismus erklärt werden können.

### 2.1.1 Die Rescorla-Wagner-Theorie

Es gibt eine Reihe konkurrierender Theorien des assoziativen Lernens. Der im Bereich kausales Lernen am häufigsten diskutierte Ansatz ist die Lernregel, die von Rescorla und Wagner (1972) vorgestellt wurde. Die von Rescorla und Wagner postulierte Lernregel geht ganz im Sinne der traditionellen Assoziations-

theorien davon aus, dass die assoziative Beziehung zwischen einem konditionierten Reiz CS und einem unkonditionierten Reiz US eine Funktion der Häufigkeit der Paarung der beiden Reize ist. Zusätzlich übernehmen Rescorla und Wagner die traditionelle Annahme, dass der Zugewinn an Assoziationsstärke proportional zur Annäherung an den durch den US festgelegten maximalen Wert abnimmt. Der Zuwachs des Assoziationsgewichts für einen CS wird als Funktion der Differenz zwischen dem aktuellen Wert und dem zu erreichenden Wert konzeptualisiert. Kognitionspsychologisch kann man diese Theorie so interpretieren, dass das Ausmaß der Veränderung eines Gewichts eine Funktion der *Überschätzung* ist. Je größer der Unterschied zwischen der durch das Assoziationsgewicht gestützten Vorhersage eines Ereignisses und der durch die Rückmeldung signalisierten richtigen Reaktion ist, desto stärker ist der Lerneffekt. Erwartet ein Hund beispielsweise am Beginn eines Lernexperiments kein Futter nach Erönen des Tons (d. h. die Assoziationsstärke zwischen den beiden Ereignissen beträgt 0), dann wird er nach Erhalten des Futters das Assoziationsgewicht in Richtung des durch die Lernerfahrung vorgegebenen Werts (z. B. 100) modifizieren. Der Zuwachs ist im Wesentlichen eine Funktion der Differenz zwischen diesen beiden Werten (100–0).

Soweit entspricht die Theorie ihren Vorläufern. Die neue Idee von Rescorla und Wagner (1972) ist nun, dass der Zugewinn an Assoziationsstärke für einen CS nicht nur eine Funktion der Vorhersage *dieser* CS ist, sondern auch abhängig ist von den assoziativen Beziehungen zwischen *anderen*, gleichzeitig präsentierten CS und dem vorherzusagenden US. Unterschiedliche CS stehen miteinander im Wettbewerb um die Vorhersage des US. Wird der US bereits durch einen anderen CS vorhergesagt, dann ist die Aufgabe erfüllt, es entsteht bei Einernen des US keine Überraschung mehr, so dass das assoziative Gewicht für keinen der aktuell wahrgenommenen CS modifiziert wird. Formal lassen sich diese Gedanken in der bekannten Lernregel von Rescorla und Wagner (1972) ausdrücken:

$$\Delta V_i = \alpha_i \beta_j (\lambda_j - \sum V_i)$$

Assoziationsgewichte werden hier mit dem Buchstaben  $V$  bezeichnet. Die Lernregel besagt, dass die Veränderung des assoziativen Gewichts zwischen einem CS $_i$  und dem US $_j$ ,  $V_i$  eine Funktion der Differenz dessen ist, was vorhergesagt werden sollte, also der durch den US $_j$  determinierten Lernasymptote  $\lambda_j$  und der Summe der assoziativen Gewichte ( $\sum V_i$ ) aller aktuell präsentierten Hinweisreize, multipliziert mit den für den US und den CS spezifischen Lernparametern  $\alpha_i$  und  $\beta_j$ . Der Lernasymptote  $\lambda_j$  wird typischerweise der Wert 100 (oder 1) bei Anwesenheit des US und der Wert 0 bei Abwesenheit zugewiesen. Häufig wird angenommen, dass die Lernraten, die die Größe der Modifikations Schritte bestimmen, größer sind, wenn der US anwesend ist, als wenn er

abwesend ist ( $\alpha_i > \beta_j$ ). In letzterem Fall werden die Assoziationsgewichte wieder reduziert, es findet also eine Extinktion (Löschung) statt. Die unterschiedlichen Lernraten werden damit begründet, dass die Anwesenheit des US ein psychologisch auffälligeres Ereignis ist als seine Abwesenheit (Rescorla & Wagner, 1972; Shanks, 1991; vgl. aber Lober & Shanks, 2000, für eine Gegenposition).

Im Abschnitt über empirische Befunde wird sich zeigen, dass diese Theorie eine Reihe von Befunden erklären kann, die in dem Bereich kausales Lernen erbracht wurden. Allgemein wird kausales Lernen hier als eine Form des Konditionierens betrachtet. Je nach Lernaufgabe übernehmen Ursachen und Effekte in der Lernsituation die Rollen von CS und US. Die Assoziationsstärke repräsentiert die Stärke der Kausalbeziehung. Kausales Lernen und kausales Vorhersagen ist also dieser Sicht zufolge lediglich ein Spezialfall der allgemeineren Kompetenz, prä-diktive Zusammenhänge zwischen Ereignissen zu lernen.

### 2.1.2 Empirische Evidenz

Die primäre Forschungsstrategie in diesem Bereich bestand zu Beginn darin zu zeigen, dass sich allgemeine Gesetzmäßigkeiten, die man in der Forschung zum klassischen und instrumentellen Konditionieren gefunden hat, wiederfinden lassen in Aufgaben, in denen es um den Erwerb von Kausalwissen geht.

In den letzten Jahrzehnten wurden viele Befunde angehäuft, die die Vorhersagen von Assoziations-theorien zu bestätigen scheinen (vgl. Allan, 1993; Shanks, 1993). So konnten z. B. Wasserman und seine Kollegen zeigen, dass Kausalurteile, analog zu den Ergebnissen der Experimente im Tierbereich, eindeutig sensitiv sind für die *Kontingenz* zwischen Ursache und Effekt (also  $\Delta P = P(c|c) - P(c|\neg c)$ ) (vgl. die Übersicht in Wasserman, 1990). Interessanterweise sagt die Rescorla-Wagner-Regel Sensitivität für Kontingenzen vorher, obwohl sie abwesende Ergebnisse nicht berücksichtigt und nur auf der Basis von Kontinguität von Ereignissen lernt. Dennoch entsprechen die Assoziationsgewichte in einer einfachen Ursache-Effekt-Lernaufgabe an der Lernasymptote der Kontingenz und nicht etwa der einfachen bedingten Wahrscheinlichkeit zwischen Ursache und Effekt (vgl. Chapman & Robbins, 1990; Cheng, 1997). Dies liegt daran, dass angenommen wird, dass bei jedem Lerdurchgang ein zusätzlicher Reiz anwesend ist, der den Kontext repräsentiert. Dieser Kontextreiz ist also sowohl vorhan-den, wenn die Ursache anwesend und abwesend ist, so dass gemäß der kom-petitiven Lernregel der Kontext mit der Ursache in Wettbewerb um die Vorhersage des Effekts geht.

Als Untersuchungsparadigma haben Wasserman und Kollegen eine einfache instrumentelle Lernprozedur gewählt, bei der die Versuchsteilnehmer nach Be-

lieben auf eine Taste drücken konnten. Variiert wurde die Kontingenz zwischen Tastendruck und einem Lichtsignal. Im Anschluss daran sollten die Versuchsteilnehmer Einschätzungen der Stärke der Kausalbeziehung auf einer Rating-skala abgeben. Wasserman und Kollegen konnten zeigen, dass die Versuchsteilnehmer sehr genau in der Lage waren, positive und negative Kontingenzen in ihrem Urteil abzubilden (Chalovich, Neunaber & Wasserman, 1985; Wasserman, Chalovich & Neunaber, 1983).

Da viele konkurrierende Theorien (z. B. Kontingenztheorien) Sensitivität für Kontingenzen vorhersagen, sind diese Befunde zwar kompatibel mit Assoziations-theorien, aber keine zwingenden Belege für ihre Gültigkeit. Interessanter ist der Befund, dass Kontingenzen, die auf häufigen Paarungen der Ereignisse beruhen, höher eingeschätzt werden als solche, die auf seltenen Paarungen beruhen (der „Dichte-Bias“). Zum Beispiel verkörpern folgende beide Situationen gleichemassen eine Kontingenz von Null zwischen Ursache (U) und Effekt (E), da die Differenz der beiden bedingten Wahrscheinlichkeiten in beiden Fällen Null beträgt:

$$A. p(c|c) = 0,8 \text{ und } p(c|\neg c) = 0,8$$

$$B. p(c|c) = 0,0 \text{ und } p(c|\neg c) = 0,0$$

Trotz identischer Kontingenzen findet man gelegentlich, dass die Versuchsteilnehmer höhere Einschätzungen in Situationen wie A abgeben als in Situationen wie B (Allloy & Abramson, 1979; Dickinson et al., 1984; Shanks, 1987), wenn-gleich in anderen Studien dieser *bias* nicht nachgewiesen werden konnte (Neunaber & Wasserman, 1986). Assoziations-theorien können das Vorliegen des *bias* nicht erklären, wenn man die Lernasymptote betrachtet, weil sie dort beide Fälle als nicht-kontingenz repräsentieren würden. Sie erklären den Effekt aber, wenn man den Lernprozess vor dem Erreichen der Lernasymptote betrachtet. Da auf Grund der kompetitiven Lernregel Kontingenzen von Null durch die Konkurrenz zwischen Ursache und Kontext gelernt werden, entstehen un-ter-schiedliche Lernsituationen. Während in der Situation A das assoziative Ge-wicht für U auf Grund der häufigen Paarung von U und E zunächst amnächst und erst allmählich wieder abfällt, wenn der Kontextreiz an Gewicht gewinnt, bleibt in Situation B das assoziative Gewicht durch das Ausbleiben einer Paarung von U und E beim initialen Gewicht von Null.

Als weiterer Beleg für die Gültigkeit assoziativer Theorien wird der Einfluss temporaler Kontinguität herangezogen. In der Regel findet man, dass die Ein-schätzung der Stärke eines Kausalzusammenhangs mit zunehmender zeitlicher Distanz von Ursache und Effekt abfällt, ähnlich wie auch Assoziationen zwischen CS und US mit zunehmender Distanz schwieriger herzustellen sind (Shanks, Pearson & Dickinson, 1989; Wasserman & Neunaber, 1986). Von Assoziations-



theorien wird dies typischerweise so modelliert, dass die Lernphase in gleich abhängige Einheiten („trials“) aufgeteilt wird. Assoziationsgewichte werden zwischen einem CS und einem US nur dann verändert, wenn sie sich innerhalb des gleichen *trial*/befinden. Andere Theorien gehen davon aus, dass das Gedächtnis für den CS mit zunehmender Zeit nachlässt (Wagner, 1981).

Das Vorliegen von *Lernkurven*, die zeigen, dass die Einschätzungen der Stärke positiver Kontingenzen inkrementell ansteigen und die negativer Kontingenzen abfallen, gilt als weiterer Beleg für assoziatives Lernen (Baker, Mercier, Vallée-Tourangeau, Frank & Pan, 1993; Chatooh et al., 1985; Lopez, Almaraz, Fernandez & Shanks, 1987). Da die statistischen Kontingenzen während des Lernverlaufs gleich bleiben, stützen diese Befunde die Deutung, dass sich während des Lernprozesses die Assoziationsgewichte gemäß der Rescorla-Wagner-Regel verändern. Interessanterweise konnte Shanks (1987) in einer Bedingung mit einer Nullkontingenz, die dem oben aufgeführten Fall A entspricht, sogar zeigen, dass die Kausalurteile zunächst ansteigen und dann erst in Richtung des Nullpunkts abfallen. Dieser Verlauf wird von der Rescorla-Wagner-Theorie vorhergesagt, die impliziert, dass das Assoziationsgewicht für den häufig mit dem US gemeinsam auftretenden CS in dem Moment abzufallen beginnt, in dem der Kontextreiz genügend Stärke erworben hat. Dieser Effekt wurde allerdings nicht immer gefunden (Baker, Bertror & Vallée-Tourangeau, 1989; Experiment 3; Waldmann, 2000).

Das vielleicht wichtigste Merkmal moderner Assoziationsstheorien ist der kompetitive Charakter der Lernregel. Es kann deshalb nicht verwundern, dass Untersuchungen zum kompetitiven Lernen den Schwerpunkt bilden, wenn es darum geht, die Gültigkeit von Assoziationsstheorien zu belegen. In der Tat wurde das Blockierungsparadigma zusammen mit anderen Paradigmen zum kompetitiven Lernen auch auf den Bereich des Erwerbs von Kausalwissen angewendet (Chapman, 1991; Chapman & Robbins, 1990; Dickinson et al., 1984; Shanks, 1985).

Ein Beispiel für das Blockierungsparadigma im Bereich kausaler Lernens stammt von Chapman und Robbins (1990). Die Lernaufgabe bestand darin, den künftigen Zustand eines fiktiven Börsenmarktes auf der Basis der gegenwärtigen Preise unterschiedlicher Aktien vorherzusagen. Der Anstieg des Preises einer Aktie entsprach dem Vorliegen einer potenziellen Ursache, blieb der Preis unverändert, dann drückte dies die Abwesenheit der potenziellen Ursache aus. Analog dazu verkörperte der Anstieg des Gesamtwertes der Börsenaktivität das Vorliegen eines Effekts, während dagegen eine monetäre Stabilität das Nicht-Vorliegen eines Effekts bedeutete. Die Lernaufgabe war in zwei Lernphasen aufgeteilt: In Phase 1 lernten die Versuchsteilnehmer, dass immer dann, wenn der Preis der Aktie P stieg, der Börsenmarkt an Wert zunahm, während immer dann, wenn der Preis der Aktie N stieg, der Markt unverändert blieb. Die Aktie P wurde also

als deterministischer Prädiktor des Marktes etabliert, während die Aktie N nicht-prädikativ war. In der entscheidenden Phase 2 wurden nun Paare von Aktien präsentiert. Dabei lernten die Versuchsteilnehmer, dass immer dann, wenn die Aktien P und B gemeinsam stiegen, der Markt an Wert zunahm, und dass, wenn die Aktien N und C gemeinsam stiegen, der Markt ebenfalls in seinem Wert zunahm. In der Testphase sollten die Versuchsteilnehmer einschätzen, wie gut die einzelnen Aktien den Markt vorherzusagen. Kritisch ist der Vergleich zwischen den Aktien B und C, die beide gleich häufig auftraten und die gleiche Korrelation zum Marktescheitern aufwiesen.

Die Rescorla-Wagner-Regel sagt vorher, dass B durch seine Paarung mit dem hochprädikativen P blockiert werden sollte, während dies bei C nicht der Fall sein sollte, das mit dem nicht-prädikativen N gepaart wurde. Blockierung ist eine Folge der additiven Komponente in der Rescorla-Wagner-Regel. In Phase 1 wird so lange trainiert, bis der prädiaktive Reiz P die Lernsymptome nahezu erreicht hat (z. B. 100). Wird in Phase 2 der neue redundante Reiz B mit P gepaart, werden gemäß der Lernregel die beiden Assoziationsgewichte für P (100) und B (0) addiert. Da die Summe (100 + 0) eine perfekte Vorhersage des Marktescheiterns erlaubt, findet keine Modifikation der Gewichte statt. Der Aufbau einer Assoziation von Aktie B wird auf Grund der Paarung mit Aktie P blockiert. Diese Vorhersage wurde von Chapman und Robbins (1990) tatsächlich bestätigt. Die Aktie C wurde als deutlich prädiaktiver eingeschätzt als die Aktie B. Neben dem Blockierungsparadigma wurde eine Reihe anderer Paradigmen des kompetitiven Lernens eingesetzt. Diese Studien belegen ebenfalls, dass es beim kausalen Lernen zur Konkurrenz von Cues kommen kann (vgl. Baker et al., 1993; Price & Yates, 1995; Wasserman, 1990).

### 2.1.3 Weiterentwicklungen von Assoziationsstheorien

Im Blockierungsparadigma wird in einer ersten Lernphase wiederholt eine Ursache mit einem Effekt gepaart, in Phase 2 sieht man dann diese prädiaktive Ursache gepaart mit einer zweiten, redundanten Ursache gefolgt von dem Effekt. In einigen Untersuchungen wurde nun die Reihenfolge der beiden Phasen umgedreht, um zu sehen, ob es auch retroaktive Blockierung gibt (vgl. Chapman, 1991; De Houwer, Beckers & Glautier, 2002; Dickinson & Burke, 1996; Larkin, Aitken & Dickinson, 1998; Wasserman & Berglan, 1998). In diesen Experimenten sieht man also in Phase 1 die beiden Ursachen gepaart und dann in Phase 2 eine der beiden Ursachen allein, gefolgt von dem Effekt. Der typische Befund ist, dass die Versuchsteilnehmer ihre Einschätzung der beiden Ursachen in Phase 2 revidieren. Während die auch in Phase 2 gezeigte Ursache als deutlich kausal für den Effekt eingeschätzt wird, revidieren die Versuchsteilnehmer die Einschätzung der nur in Phase 1 gezeigten Ursache nach unten. Es findet also

eine retrospektive Blockierung start, auch wenn der Effekt meist kleiner ausfällt als im Standardparadigma. Dieser so genannte *retrospektive Revaluationseffekt* lässt sich nicht mit der Rescorla-Wagner-Regel erklären. Da die Regel davon ausgeht, dass nichts über abwesende Reize gelernt wird, wird das in Phase 1 etablierte Assoziationsgewicht der Ursache, die in Phase 2 nicht mehr präsentiert wird, nach Abschluss der ersten Lernphase nicht verändert.

Eine Reihe modifizierter Assoziationsstheorien wurde entwickelt, die retrospektive Revaluation erklären können. Van Hamme und Wasserman (1994) schlugen eine Variante der Rescorla-Wagner-Regel vor, bei der angenommen wurde, dass assoziative Gewichte auch für abwesende Cues revidiert werden, sofern sie auf Grund der Instruktion oder früherer Paarung mit anderen Reizen erwartet werden. So erzeugt die Paarung der beiden Ursachen in Phase 1 eine Assoziation zwischen diesen beiden Cues. In Phase 2 wird daher erneut die Anwesenheit beider Cues erwartet, was aber nicht der Fall ist. Das Modell sieht vor, dass der Lernparameter  $\alpha$  einen positiven Wert bei anwesenden CS und einen negativen bei abwesenden, aber erwarteten CS einnimmt. Ein verwandtes Modell, das auf dem Standard Operating Procedures (SOP)-Modell von Wagner (1981) aufbaut, wurde von Dickinson und Burke (1996) vorgelagert. Dieses Modell geht davon aus, dass in der ersten Phase eine Assoziation zwischen den beiden gepaarten CS aufgebaut wird. Latkin et al. (1998) testeten diese Annahme experimentell und konnten zeigen, dass die Stärke des Revaluationseffekts von der Stärke der Assoziation zwischen den CS in Phase 1 abhängt.

Die Frage nach Assoziationen zwischen Cues beschäftigt auch eine andere Forschungsrichtung. Die Rescorla-Wagner-Regel und verwandte Modelle gehen davon aus, dass einzelne Cues (z. B. Ursachen) mit den Effekten assoziiert werden. Einige Studien konnten aber zeigen, dass die Versuchsteilnehmer auch Konfigurationen von Cues als eigenständige Reize betrachten können (z. B. Shanks, Charles, Darby & Azini, 1998). So konnten Williams, Sagness und McPhee (1994) zeigen, dass der Blockierungseffekt im Standardparadigma verschwindet, wenn die Versuchsteilnehmer die gepaarten Ursachen als konfigurale Einheiten sahen und nicht als elementare Cues. Eine Reihe neuerer Theorien hat sich deshalb der Frage gewidmet, wie man konfigurales Lernen modellieren kann (z. B. Pearce, 1994; Wagner & Brandon, 2001).

Die meisten Assoziationsstheorien modellieren die Modifikation von Assoziationsgewichten während des Lernens. Im Gegensatz dazu haben Miller und Kelleys (vgl. Miller & Marzel, 1988) eine Theorie vorgelagert, die auf den Zeitpunkt des Testens fokussiert. Gemäß der von diesen Autoren vorgeschlagenen *Komparativtheorie* werden beim Lernen einfache Assoziationen mittels eines nicht-komparativen, assoziativen Lernmechanismus zwischen den Reizen aufgebaut. Das heißt, es werden in Abhängigkeit der Häufigkeit raum-zeitlicher Paarung alle präsentier-

ten Reize (ob CS oder US) miteinander assoziiert. Beim Testen wird dann die Assoziation zwischen dem CS, nach dem gefragt wird, und dem US verglichen. Sagen diese anderen Reize den US ebenfalls vorher, dann führt dies dazu, dass der Zielreiz eher nicht genutzt wird. Dieses Modell kann den Blockierungseffekt in beiden Varianten (vorwärts und rückwärts) erklären. Blaisdell, Bischof, Guther und Miller (1998) haben ein erweitertes Komparativmodell vorgelagert, das noch komplexere Vergleiche erlaubt.

#### 2.1.4 Theoretische Kritik an Assoziationsstheorien

Der letzte Abschnitt hat deutlich gemacht, dass Assoziationsstheorien weiterentwickelt wurden, um kritische Befunde erklären zu können. Dennoch gibt es eine Reihe von Eigenschaften, die diese Theorien als Modelle für Kausalität kritisch erscheinen lässt. Ein zentrales Merkmal ist, dass diese Modelle über kein Gedächtnis für Ereignisse verfügen. Assoziationsgewichte spiegeln den aktuellen Stand wider; wie es aber zu dem jeweiligen Gewicht kann, bleibt verborgen. Selbst die Ansätze, die retrospektive Revaluation ermöglichen, modifizieren die Einschätzung nicht deshalb, weil sie Ereignisse aus der Vergangenheit heranziehen und neu bewerten. Die postulierten Mechanismen gehen vielmehr davon aus, dass Lernen in aktuellen und bei künftigen Lerndurchgängen auf Grund der Erwartung von aktuell abwesenden Cues modifiziert wird. Diese Modelle haben Probleme mit Befunden, die echte Retrospektion belegen. So konnten Waldmann (2000) und De Houwer (2002) zeigen, dass der Blockierungseffekt davon abhängt, ob die Versuchsteilnehmer annehmen, dass der redundante Cue in der Phase, in der er nicht gezeigt wurde, abwesend war oder lediglich nicht beobachtet wurde. Eine Blockierung resultierte nur dann, wenn die Teilnehmer davon ausgingen, dass der redundante Cue in dieser Phase wirklich abwesend war. Ein weiterer kritischer Befund stammt von Macho und Burkart (2002), die zeigen konnten, dass Versuchsteilnehmer komplexe Revaluationen vornehmen können, bei denen Cues mit anderen Cues verglichen wurden, deren Einschätzung von einer weiteren Revaluation abhängt (rekursive Revaluation) (vgl. auch De Houwer & Beckers, 2002a, b).

Eine zweite Klasse von Einwänden beschäftigt sich mit der Frage, ob Assoziationsstheorien grundlegenden Eigenschaften von Kausalität gerecht werden. Assoziationsstheorien sind formal vergleichbar mit dem statistischen Verfahren der multiplen Regression, bei dem mehrere Prädiktoren mit einem vorherzusagenden Kriterium in Verbindung gebracht werden. Welche kausale Rolle Prädiktoren und Kriterien spielen, geht dabei in die Modellierung nicht ein. Prädiktoren können also Ursachen, Effekte oder Mischungen der beiden Klassen sein. Entscheidend ist lediglich, welche Ereignisse als Cues zur Vorhersage verwendet werden. Damit drücken sie aber ein wesentliches Merkmal von Kausalzusammenhängen

nicht aus, nämlich deren Gerichtetheit (vgl. Waldmann, 1996). Ursachen geometrischen Effekte aber nicht wie versa. Assoziationstheorien können mithin auch keine komplexeren Kausalmodelle repräsentieren (siehe Abschnitt 2.5). Eine Debatte hat sich in den letzten Jahren deshalb der Frage gewidmet, ob Versuchsteilnehmer beim Lernen von Kausalmodellen die Gerichtetheit von Kausalität berücksichtigen oder nicht. Während Waldmann und Kollegen (Waldmann, 2000, 2001; Waldmann & Hagemyer, 2001; Waldmann & Holyoak, 1992; Waldmann, Holyoak & Frerianne, 1995; Waldmann & Wälschen, 2005) Belege dafür vorgelegt haben, dass Versuchsteilnehmer sensitiv sind für den Aspekt der Kausalrichtung, gibt es auch Befunde, die die assoziations-theoretische Sicht zu bestätigen scheinen (Cobos, Lopez, Cano, Almaraz & Shanks, 2002; Shanks & Lopez, 1996).

Die Nutzung von Wissen über die Kausalrichtung ist nur ein Beispiel, wie abstraktes Wissen über Kausalität den Lernprozess beeinflussen kann. Hagemyer und Waldmann (2002) haben sich mit der von Assoziationstheorien vertretenen Annahme befasst, dass in der Regel nur raum-zeitlich nahe beieinander liegende Ereignisse miteinander assoziiert werden (Kontingenzannahme). Sie konnten in einer Reihe von Experimenten zeigen, dass die Versuchsteilnehmer Vorannahmen über die Dauer kausaler Mechanismen bei der Bestimmung der kausal verbundenen Ereignisse nutzen. Je nachdem, ob nahe oder ferne Ereignisse in einem Ereignisstrom miteinander verknüpft wurden, kam es zu dramatisch unterschiedlichen Einschätzungen der kausalen Stärke.

Schließlich hat Cheng (1997) argumentiert, dass Assoziationstheorien die Beziehung von Kovariation und kausaler Power falsch konzeptualisieren (vgl. Abschnitt 2.2 für eine Darstellung ihrer Theorie). Ihrer Sicht zufolge setzen Assoziationstheorien kausale Power irrtümlicherweise gleich mit Kovariationen. Möchte man beispielsweise untersuchen, ob ein Medikament gegen Allergien hilft, dann wird man dies nur in volleren Umlauf in den Daten sehen, wenn man nur Probanden untersucht, die Allergien haben. UnTERSUCHT man aber gesunde Patienten, dann wird sich in den Daten auf Grund des Bodeneffekts kein Effekt zeigen (Kontingenz von 0). Umgekehrt kann man beispielsweise den Effekt von Rauchern auf die Atmung nicht erkennen, wenn man nur Patienten untersucht, die auf Grund anderer Ursachen bereits an Atembeschwerden leiden (Deckeneffekt). Assoziationstheorien würden in diesen beiden Situationen von der Nullkontingenz darauf schließen, dass es keinen kausalen Zusammenhang gäbe. Richtig wäre es aber, die Basisrate des Effekts (also der Anteil an Fällen, der auf andere, unbekannte Ursachen zurückzuführen ist) zu berücksichtigen. Lober und Shanks (2000) stimmen zu, dass dies normativer richtig wäre. Ihre Experimente deuten aber darauf hin, dass Versuchsteilnehmer nicht normativ urteilen, sondern eher im Sinne der Rescorla-Wagner-Theorie. Demgegenüber haben Buchner, Cheng und Clifford (2003) gezeigt, dass die Urteile von Probanden konsistent

sind mit den Vorhersagen in Chengs Theorie (vgl. Abschnitt 2.2 für weitere Details). Auch hier bemüht sich derzeit eine Reihe von Forschern darum, die Randbedingungen zu klären, die Versuchsteilnehmer zu normativen oder nicht-normativen Urteilen verleiten.

## 2.2 Regelbasierte Ansätze

Regelbasierte Ansätze untersuchen, wie Informationen über das Vorliegen oder Fehlen eines Effekts bei An- oder Abwesenheit der Ursache das Kausalurteil beeinflussen. Sie gehen in der Tradition Humes davon aus, dass der statistische Zusammenhang zwischen Ursache und Wirkung der wesentliche Indikator für das Vorliegen einer Kausalrelation ist. Die einzelnen Theorien unterscheiden sich erstens darin, welche Regeln sie als grundlegend annehmen, und zweitens darin, ob sie ein normatives Modell kausalen Urteilens aufstellen oder ob sie menschliches Urteilen lediglich beschreiben. Der erste der hier näher beschriebenen Ansätze, der varianzanalytische Ansatz nach Kelley (1973, 1967), hat den Anspruch, normativ zu sein und modelliert die Kausalattribution nach dem Vorbild der Varianzanalyse in der Statistik. Die kontingenzbasierten Ansätze beziehen sich auf probabilistische Kausalitätskonzepte in der Philosophie und erheben ebenfalls den Anspruch normativ zu sein. Bei diesen muss zwischen einfachen und komplexen Ansätzen differenziert werden: Ersterer beziehen sich nur auf einfache Ursache-Effektbeziehungen, während Letztere davon ausgehen, dass in der Welt stets mehrere Ursachen gleichzeitig einen Effekt beeinflussen. Nach der Diskussion dieser grundlegenden Theorien folgen drei Ansätze, welche die Kontingenztheorien weiterentwickeln und verschiedene Schwächen der herkömmlichen Theorien nicht teilen. Hierzu gehört das probabilistische Konstruktmodell nach Cheng und Novick (1990, 1992), die Theorie der kausalen Power nach Cheng (1997) und die Theorie der Evidenzevaluation nach Whittle (2003).

### 2.2.1 Der varianzanalytische Ansatz

*Theorie.* Der varianzanalytische Ansatz (Kelley, 1967, 1972, 1973) baut auf früheren Ansätzen zur Kausalattribution (Heider, 1958) auf. Nach diesem Ansatz versucht ein Beobachter in der Rolle eines naiven Wissenschaftlers Ereignissen, v. a. menschliche Verhaltensweisen, zu erklären, indem er die zugehörigen Ursachen bestimmt. Kelley geht davon aus, dass die Kausalattribution auf dem Kovariationsprinzip beruht. Damit ein Faktor als Ursache für ein Ereignis relevant sein kann, muss er mit diesem kovariieren. Das Kovariationsprinzip lautet in den Worten Kelleys (1967, S. 194): „The effect is attributed to that condition which is present, when the effect is present and absent, when the effect is

absent“. Die Formulierung des Prinzips macht deutlich, dass hier von deterministischen Zusammenhängen von Ursache und Wirkung ausgegangen wird. Der Theorie zufolge machen sich Personen die beobachtbaren Kovariationen zueinander zu Nutze. Die Personen führen implizit eine Varianzanalyse für mögliche Ursachenfaktoren durch und bestimmen so Haupteffekte und Interaktionen der Ursachen in Bezug auf den Effekt.

Kelley hat seine Theorie speziell für die Kausalerklärung von menschlichen Verhaltensweisen entwickelt, so dass er nur der Klassen möglicher Ursachen von Verhaltensweisen annimmt. Die Ursache des Verhaltens kann in der Person, in den gegebenen Stimuli oder in der Situation (dem spezifischen Zeitpunkt) liegen. Die zu den drei Ursachendimensionen zugehörigen Kovariationsinformationen sind Konsens, d. h. inwieweit auch andere Personen in derselben Situation bei gleichen Stimuli sich gleich verhalten, die Distinktheit, d. h. inwieweit die Zielperson auch bei anderen Stimuli das Verhalten zeigt, und die Konsistenz, d. h. inwieweit die Zielperson auch zu anderen Zeitpunkten das Verhalten zeigt. Aus dem Kovariationsprinzip folgt: dass das Verhalten bei niedrigem Konsens, niedriger Distinktheit und hoher Konsistenz auf die Person attribuiert werden sollte. Eine Attribution auf den Stimulus sollte bei hohem Konsens, hoher Distinktheit und hoher Konsistenz und eine Attribution auf die Situation sollte bei niedrigem Konsens, hoher Distinktheit und niedriger Konsistenz erfolgen.

Zusätzlich zu dem allgemeinen varianzanalytischen Ansatz entwickelte Kelley das Konzept der kausalen Schemata, die im Alltag ebenfalls bei der Ursachenerklärung Anwendung finden (Kelley, 1972). So behauptet Kelley, dass Personen in der Regel davon ausgehen, dass jede von verschiedenen Ursachen allein hinreichend ist, den Effekt hervorzurufen (Schema multipler hinreichender Ursachen). In diesem Fall sollte das Wissen über das Vorliegen einer der Ursachen dazu führen, dass das Vorliegen anderer Ursachen für nicht mehr kausal relevant erachtet wird (Abwertungsprinzip). Bei extremen oder besonderen Ereignissen wenden Personen aber ein anderes Schema, das Schema multipler notwendiger Ursachen, an. Dieses besagt, dass mehrere Ursachen zugleich gegeben sein müssen, damit der zu erklärende Effekt überhaupt eintritt. In einem solchen Fall sollte das Vorliegen anderer Ursachen nicht wie beim Schema multipler hinreichender Ursachen abgewertet, sondern aufgewertet werden.

*Empirische Evidenz.* In Untersuchungen zur Kausalattribution wurden Probanden mit bestimmten Verhaltensweisen einzeln oder Personen in einer bestimmten Situation konfrontiert, beispielsweise damit, dass eine fiktive Jane zu einer bestimmten Gelegenheit Spaß beim Abwaschen hatte. Zusätzlich bekamen die Versuchsteilnehmer Informationen zu Konsens (z. B. „Niemand sonst hatte Spaß“),

Distinktheit (z. B. „Sie hat auch bei anderen Haushaltstätigkeiten Spaß“) und Konsistenz (z. B. „Abwaschen macht ihr immer wieder Spaß“). Die Aufgabe der Probanden war es zu bestimmen, ob das Verhalten durch etwas in der Person, dem Stimulus oder der Situation (dem spezifischen Zeitpunkt) hervorgerufen wurde.

In einer Reihe von Studien wurden die Vorhersagen des varianzanalytischen Ansatzes untersucht (McArthur, 1972; 1976; Orvis, Cunningham & Kelley, 1975; siehe auch Kelley & Michela, 1980, für eine Übersicht). Die Ergebnisse bestätigten die Vorhersagen in Bezug auf die drei oben ausgeführten spezifischen Informationsmuster. Für andere aus dem Kovariationsprinzip ableitbare Vorhersagen konnte allerdings keine Evidenz gefunden werden. Zwei abweichende Befunde erwiesen sich dabei als so wichtig, dass sie als eigenständige Biases Eingang in die Literatur fanden: Der Bias, Konsensusinformation zu vernachlässigen, und der Bias, Verhalten tendenziell eher auf die Person zu attribuieren. Der erste Bias konnte bereits in den ersten empirischen Untersuchungen zum varianzanalytischen Ansatz nachgewiesen werden (McArthur, 1972) und wurde später vielfach repliziert. Die Vernachlässigung der Konsensusinformation kann beispielsweise dazu führen, dass wir die Wahrscheinlichkeit, dass andere Personen ähnliche Merkmale wie wir haben, überschätzen. Der zweite Bias wurde ebenfalls mehrfach belegt (Jaspars, Hewstone & Fiedler, 1983; McArthur, 1972; Orvis et al., 1975). So zeigte sich beispielsweise, dass auch, wenn keinerlei Kovariation des Ereignisses zu den drei Ursachendimensionen vorlag, d. h. bei hohem Konsens, niedriger Distinktheit und hoher Konsistenz, Personen dazu tendierten, das Ereignis auf die Person zu attribuieren (McArthur, 1972; Orvis et al., 1975). Personen tendieren außerdem dazu, Verhalten eher auf die Absichten oder die Persönlichkeitsmerkmale einer anderen Person zu attribuieren als auf situative Merkmale (fundamentaler Attributionsfehler). Auch kommt es zu selbstwertdienlichen Verzerrungen, beispielsweise werden Erfolge eher der eigenen Person zugeschrieben, während die Ursache für Misserfolge eher in der Situation gesucht wird. Die Lage verkompliziert sich zusätzlich dadurch, dass diese Biases nicht unter allen Bedingungen auftreten (vgl. Cheng & Novick, 1990). In den letzten Jahren mehrten sich auch Befunde, die zeigen, dass manche Biases (etwa der fundamentale Attributionsfehler) vor allem im westlichen Kulturkreis auftreten, nicht aber im östlichen (vgl. Nisbett, Peng, Chohi & Norenzayan, 2001). Die Befundlage für die Anwendung kausaler Schemata ist ebenfalls uneinheitlich und unterstützt nicht durchgängig den Ansatz von Kelley (vgl. Fiedler, 1982, für einen Überblick).

*Theoretische Kritik am varianzanalytischen Ansatz.* Ein theoretisch gewichtiger Einwand gegen den varianzanalytischen Ansatz war, dass auf der Grundlage von Konsens-, Distinktheits- und Konsistenzinformation keine vollständige Varianzanalyse durchgeführt werden kann, da notwendige Informationen für eine solche Analyse fehlen und daher von den Probanden zusätzliche Annahmen ge-

macht werden müssen (Cheng & Novick, 1990; Forstling, 1989; Hilton, 1988, 1990). Dies wird deutlich, wenn man sich das oben genannte Beispiel in Form bedingter Wahrscheinlichkeiten vergegenwärtigt: Jane hat an einem bestimmten Tag Spaß beim Abwaschen ( $P(\text{Spaß} | \text{Jane, Abwaschen, TagX}) = 1$ ), niemand sonst hat Spaß ( $P(\text{Spaß} | \text{Jane, Abwaschen, TagX}) = 0$ ), Jane hat auch Spaß an anderen Hausarbeitstätigkeiten ( $P(\text{Spaß} | \text{Jane, Abwaschen, TagX}) = 1$ ) und Jane hat immer wieder Spaß beim Abwaschen ( $P(\text{Spaß} | \text{Jane, Abwaschen, TagX}) = 1$ ). Wie die Darstellung zeigt, fehlt eine ganze Reihe möglicher Bedingungsombinationen, so beispielsweise die Information, wie wahrscheinlich es ist, dass jemandem an einem beliebigen Tag Spaß beim Abwaschen hat ( $P(\text{Spaß} | \text{Jane, Abwaschen, TagX})$  oder die Information, dass jemand zu einem beliebigen Zeitpunkt Spaß bei dem hat, was er oder sie tut ( $P(\text{Spaß} | \text{Jane, Abwaschen, TagX})$ ). Die Unvollständigkeit der Information wird als ein möglicher Grund für die abweichenden Befunde angesehen (vgl. den Abschnitt zum Probabilistischen Kontrastmodell).

Neben der Unvollständigkeit der innerhalb des Ansatzes verwendeten Informationen wurde auch die Beschränkung auf deterministische Zusammenhänge kritisiert. Viele Zusammenhänge in der Welt sind probabilistischer Natur. Auch die im Rahmen einer Varianzanalyse untersuchten Zusammenhänge müssen nicht deterministisch sein.

Ein dritter wesentlicher Kritikpunkt ist, dass sich der Ansatz nur auf die Kausalattribution bezieht und damit den wichtigen Bereich des Erwerbs von neuem Wissen über allgemeine Kausalzusammenhänge vernachlässigt. Dies ist verdeutlicht, da auch hier das Kovariationsprinzip angewendet werden kann (vgl. Abschnitt zu assoziativen Ansätzen und kontingenzbasierten Ansätzen).

Die inhaltliche Beschränkung auf Verhaltensweisen und die zugehörigen drei Ursachendimensionen ist ein weiterer Kritikpunkt. Dies reduziert die Anwendbarkeit des Ansatzes in unnötiger Weise. Das Kovariationsprinzip ist in seiner ursprünglichen Formulierung allgemein gültig und kann daher auf alle Arten von Ursache-Wirkungszusammenhängen angewendet werden.

Nichtsdetrotor war der varianzanalytische Ansatz der Kausalattribution der vielleicht wichtigste Ausgangspunkt für die sozialpsychologische Forschung zur Kausalattribution von Verhaltensweisen sowie sozial und persönlich relevanten Ereignissen. So konnte gezeigt werden, dass Kausalattributionen vor allem dann vorgenommen werden, wenn unerwartete Ereignisse auftreten, gesetzte Ziele nicht erreicht werden oder starke Emotionen wie Traurigkeit oder Ärger auftreten (vgl. Weiner, 1985). Als wichtige Ursachendimensionen stellte sich die Unterscheidung von internen (personnanmenten) und externen (umweltbezogenen) Ursachen, von stabilen versus instabilen und kontrollierbaren versus nicht durch

die Person kontrollierbaren Ursachen heraus (Weiner, 1986). Eine ausführliche Darstellung aktueller sozialpsychologischer Attributionstheorien würde den Rahmen dieses Kapitels sprengen. Der Leser sei daher auf die Einführung von Friedman und Hewstone (2002) und die Übersichtsarbeiten von Gilbert (1995) und Forstling (2001) verwiesen.

### 2.2.2 Kontingenztheoretische Ansätze

Während der varianzanalytische Ansatz genauso wie viele philosophische Ansätze davon ausgeht, dass Kausalzusammenhänge deterministisch sind, nehmen kontingenztheoretische Ansätze an, dass Kausalzusammenhänge probabilistisch sein können. Diese Ansätze bauen damit auf probabilistischen Kausalitätstheorien in der Philosophie auf (Cartwright, 1983, 1989; Ellis, 1991; Salmon, 1980, 1984; Suppes, 1970, 1984). Gemäß diesen Theorien ist ein Ereignis C dann eine Ursache eines anderen Ereignisses E, wenn die Wahrscheinlichkeit des Effektereignisses E bei Vorliegen des Ursachereignisses C ungleich der Wahrscheinlichkeit des Effektereignisses E bei Abwesenheit der Ursache ereignisses C ist,  $P(E|C) \neq P(E|-C)$ . Ist die Wahrscheinlichkeit des Effektes bei Anwesenheit größer als bei Abwesenheit der Ursache, so wirkt die Ursache generativ, sie ruft also den Effekt hervor. Ist die Wahrscheinlichkeit bei Abwesenheit größer als bei Anwesenheit der Ursache, so wirkt die Ursache inhibitorisch, sie unterdrückt also den Effekt. Der Kausalzusammenhang zwischen den Ereignissen wird durch die Differenz der beiden Wahrscheinlichkeiten bestimmt. Diese Differenz wird als Kontingenz bezeichnet und ist definiert als  $\Delta P = P(E|C) - P(E|-C)$ .

*Einfache Kontingenztheorien.* Diese Theorien gehen davon aus, dass Personen beim Erwerb von Kausalwissen bzw. bei der Urteilsbildung über die Stärke eines Kausalzusammenhangs die Kontingenz zwischen der vermeintlichen Ursache und dem in Frage stehenden Effekt ermitteln. Personen werden also als intuitive Statistiker angesehen, welche implizit Kontingenzen auf der Grundlage der verfügbaren Informationen bestimmen. Die Informationen über den Zusammenhang zwischen einer Ursache C und einem Effekt E werden in der Regel in Form einer Vier-Felder Kontingenztafel dargestellt, wobei die einzelnen Zellen mit Buchstaben von A bis D gekennzeichnet werden (siehe Abbildung 1).

	Effekt gegeben	Effekt abwesend
Ursache gegeben	A	B
Ursache abwesend	C	D

Abbildung 1  
4-Felder Kontingenztafel

Die Kontingenz von Ursache und Wirkung ergibt sich dann als  $AD^2 = (A/(A+B)) - (C/(C+D))$  (Allan & Jenkins, 1983; Ward & Jenkins, 1965).

*Empirische Evidenz.* Die Literatur zu Kontingenzschätzungen ist sehr umfangreich und erstreckt sich nunmehr über fast ein halbes Jahrhundert. Vielfältige Paradigmen wurden eingesetzt. So wurde variiert, in welcher Form die Informationen dargeboten wurden, entweder als summarische Tabellen, in Form von Falllisten oder als Sequenz von Einzelfällen in Form eines „trial-by-trial“-Lernens. Variiert wurden die Besetzungen der einzelnen Zellen. Analysten wurden die Einschätzungen des Zusammenhangs durch die Probanden auf zwei verschiedenen möglichen Regeln der Integration der Zellen berechnet (z. B. Ward & Jenkins, 1965), oder die Regeln wurden aus den Einschätzungen jeder Versuchsperson bei einer Vielzahl von Aufgaben rekonstruiert (Shallice, 1983; Shallice & Tucker, 1980).

Die Ergebnisse dieser Studien sind uneinheitlich. Eine Reihe von Studien berichtet positive Evidenzen für die Verwendung von  $AD^2$  (Ward & Jenkins, 1965; Wasserman et al., 1983). Andere Studien nennen die so genannte „Summe der Diagonalen“-Regel als die häufigste oder eine der häufigsten Regeln (Atkes & Harkness, 1983; Allan & Jenkins, 1983; Shallice, 1983; Shallice & Tucker, 1980). Bei dieser auch  $AD$  genannten Regel wird die Summe der für einen Zusammenhang sprechenden Fälle in den Zellen A und D von der Summe der gegen einen Zusammenhang sprechenden Fälle in den Zellen C und B abgezogen,  $AD = (A + D) - (C + B)$ . Weitere empirisch gefundene Regeln waren die Differenzbildung der Zellen A und B, der Zellen A und C oder auch nur die Verwendung der A-Zelle. Die Ergebnisse insgesamt deuten darauf hin, dass die Licht dargestellt werden und die Aufgabe eine Berücksichtigung aller gegebenen Informationen nahe legt. Die Kontingenz wird auch genutzt, wenn die Probanden zur Bedeutung der gestellten Fragen vgl. Beryl-Marion, 1982). Vereinfachende Regeln werden vor allem dann eingesetzt, wenn die Aufgabe komplex ist (Kao & Wasserman, 1993; Shallice & Mims, 1982; Ward & Jenkins, 1965).

Anderer Studien haben sich mit der Gewichtung der einzelnen Zellen für die kausale Urteilbildung auseinandergesetzt. Hierbei zeigten sich überraschend übereinstimmende Ergebnisse. Die A-Zelle wird am stärksten und die D-Zelle am wenigsten gewichtet. Insgesamt ergab sich folgende Abfolge:  $A > B > C > D$  (Kao & Wasserman, 1993; Schustack & Sternberg, 1981; Wasserman, Donner & Kao, 1990). Ein weiterer wichtiger Befund war, dass bei konstantem  $AD^2$  die Einschätzungen mit der Wahrscheinlichkeit des Effektes ansteigen. Dieser Effekt kann von einfachen Kontingenztheorien nicht erklärt werden (vgl. Allan, 1993).

*Theoretische Kritik am einfachen kontingenztheoretischen Ansatz.* Obwohl viele Forscher die Kontingenz als normatives Maß für den Kausalzusammenhang zweier Ereignisse ansehen, lässt sich nachweisen, dass die Kontingenzberechnung unter bestimmten Bedingungen zu Fehleinschätzungen führt (vgl. Cheng, 1993, 1997; Cheng & Novick, 1992). Ein erster Fall ist der, in dem der Effekt durch eine andere (unbekannte oder un beobachtete) Ursache immer hervorgerufen wird. In diesem Fall ist die Wahrscheinlichkeit des Effektes bei An- und Abwesenheit der zu untersuchenden Ursache immer gleich Eins und die Kontingenz damit gleich Null ( $AD^2 = P(e|c) - P(e|\bar{c}) = 1 - 1 = 0$ ). Dennoch würden wir intuitiv nicht sagen wollen, dass die zu untersuchende Ursache keinen Einfluss hat, sondern dass wir unter diesen Bedingungen keine Beurteilung vornehmen können. Wenn wir beispielsweise wissen wollen, ob ein bestimmtes Medikament Allergien erzeugt, dann würden wir die Frage nicht mit „Nein beantworten“, wenn wir zufälligerweise nur solche Patienten mit dem Medikament behandelt haben, die bereits vorher schon an der Allergie litten.

Ein zweiter Fall betrifft das Vorhandensein notwendiger Randbedingungen, beispielsweise das Vorhandensein von Sauerstoff für die Entstehung eines Feuers. Um die Wirkung von Funkenflug auf die Entstehung von Feuer untersuchen zu können, ist es notwendig, dass Sauerstoff vorhanden ist. Ansonsten müssten wir schließen, dass Funkenflug keinerlei Einfluss hat. Solche Rahmenbedingungen finden aber keinen Eingang in einfache Kontingenzformeln.

Ein weiterer wichtiger theoretischer Einwand gegen einfache Kontingenztheorien ist, dass diese nicht in der Lage sind, zwischen echten (genannten) Kausalrelationen und scheinbaren („spurious“) Relationen zu unterscheiden. Scheinbare Kausalrelationen sind statistisch zuverlässige Zusammenhänge zwischen Ereignissen, die aber nicht auf dem Vorhandensein eines direkten kausalen Einflusses beruhen. Ein beliebigeres Beispiel hierfür ist der Zusammenhang zwischen dem Stand eines Barometers und einem Gewitter. Die positive Relation der beiden beruht darauf, dass beide vom Luftdruck kausal abhängen, selbst aber nicht direkt kausal verbunden sind. Aus einer einfachen Kontingenz allein ist aber nicht zu ersehen, ob der Zusammenhang kausal ist oder nicht (vgl. Cheng & Novick, 1992; Waldmann & Hagnayer, 2001).

### 2.2.3 Komplexer, multifaktorieller Kontingenztheorien

Komplexe, multifaktorielle Kontingenztheorien gehen davon aus, dass stets mehrere Faktoren kausal relevant für ein Ereignis sind. Diese zusätzlichen Faktoren können dazu führen, dass die zwischen einer bestimmten Ursache und dem Effekteffekt beobachtete Kontingenz nicht den tatsächlichen Kausaleinfluss widerspiegelt (vgl. Kritik einfacher Kontingenztheorien). Eine Reihe von Philoso-

phen (Carwright, 1983; Ellis, 1991; Salmon, 1980) und kognitiver Psychologen (Cheng, 1993; Spellman, 1996a/b; Spellman, Price & Logan, 2001; Waldmann, 1996; Waldmann & Hagemayer, 2001; Waldmann & Holyoak, 1992) hat daher vorgeschlagen, weitere kausal relevante Faktoren zu berücksichtigen und an Stelle der unbedingten bedingte Kontingenzen zu verwenden. Bedingte Kontingenz bedeutet, dass bei der Bestimmung der Kontingenz zwischen einer Ursache C und einem Effekt E das Vorliegen der anderen kausal relevanten Faktoren  $K_1$  bis  $K_n$  konstant gehalten wird  $\Delta P = P(E|C, K_1, \dots, K_n) - P(E|C, K_1, \dots, K_n)$ . Durch die Konditionalisierung auf andere kausal relevante Faktoren können mögliche Interaktionen mit anderen Faktoren entdeckt werden. Um beispielsweise die Hypothese, dass Rauchen (C) Herzkrankheiten (E) hervorruft, zu testen, sollte mit berücksichtigt werden, dass Raucher häufig auch andere gesundheitsbedingliche Angewohnheiten wie hohen Kaffeekonsum (K) haben. Bei der Testung sollte daher der Einfluss des Kaffeekonsums kontrolliert werden. Die Kontingenz von Rauchen und Herzkrankheiten sollte deshalb getrennt in der Gruppe der Kaffeetrinker,  $\Delta P = P(E|C, K) - P(E|C, K)$ , und der Nicht-Kaffeetrinker,  $\Delta P = P(E|C, \neg K) - P(E|C, \neg K)$ , ermittelt werden. Wenn das Rauchen ein unabhängig wirkender Ursachenfaktor für Herzkrankheiten ist, sollte sich in beiden Gruppen eine positive Kontingenz zeigen.

*Empirische Evidenz:* Mehrere Studien in den letzten Jahren haben gezeigt, dass Personen zusätzlich vorliegende Faktoren bei der Bestimmung von Kausalrelationen berücksichtigen (Baker et al., 1993; Price & Yates, 1993; Spellman, 1996a/b; Spellman et al., 2001; Waldmann & Hagemayer, 2001). In einem Experiment von Spellman (1996a) bekamen Probanden beispielsweise die Aufgabe, den kausalen Einfluss zweier Flüssigkeiten (vernünftlichen Düngers) auf das Blühen von Pflanzen zu untersuchen. Varietäten wurden in den Experimenten die bedingten und unbedingten Kontingenzen der beiden Flüssigkeiten und des Blühens. Die Ergebnisse zeigen, dass die Probanden bei ihren Einschätzungen des Kausalflusses auf das Vorhandensein des jeweils anderen Düngers konditionalisieren, also die bedingten Kontingenzen nutzen.

Multi-faktorielle Kontingenztheorien sind auch in der Lage, den Blockierungseffekt zu erklären (vgl. Abschnitt 2.1 zu den Assoziations-theorien). In der ersten Phase der Blockierungsparadigma ist die unbedingte Kontingenz der einen Ursache auf den Effekt gleich Eins. In der zweiten Phase wird diese erste Ursache mit einer zweiten Ursache deterministisch gepaart und der Effekt tritt stets auf, wenn beide Ursachen deterministisch gepaart und der Effekt tritt stets auf, wenn beide Ursachen gegeben sind. So lernen beispielsweise in einem Experiment von Waldmann (2000) die Versuchsteilnehmer in Phase 1, dass eine Substanz in der Nahrung eine deterministische Ursache einer Blutkrankheit ist. In Phase 2 wurde diese Substanz stets mit einer zweiten, redundanten Substanz gepaart. Beides zusammen hatte ebenfalls die Krankheit zur Folge. Sollten die Versuchsteilnehmer die kausalen Beziehungen einschätzen, dann sahen sie am

Ende die erste Substanz als Ursache, bei der zweiten waren sie aber unsicher, obwohl diese für sich genommen ebenfalls eine unbedingte Kontingenz von Eins hatte. Einräthige Kontingenztheorien können den Unterschied also nicht erklären (vgl. auch Waldmann, 1996; Waldmann & Holyoak, 1992). Multi-faktorielle Kontingenztheorien können dies aber. Gemäß dieser Theorie muss bei der Bestimmung der Kontingenz der zweiten Ursache auf das Vorliegen der ersten Ursache konditionalisiert werden. Dies ist aber nicht vollständig möglich, da die zweite Ursache niemals alleine vorliegt. Die Folge ist, dass die Probanden sich über den kausalen Status der zweiten Ursache unsicher sein sollten (vgl. auch Cheng, 1997; Waldmann & Hagemayer, 2001). Die Einschätzungen der Probanden in einer Vielzahl von Studien deuten darauf hin, dass diese Interpretation des Blockierungseffekts richtig ist. Die Urteile lagen durchweg im mittleren Bereich der verschiedenen Ratingskalen und nicht am unteren Ende. Das heißt die Probanden sagten nicht, dass die zweite Ursache keinen Einfluss hatte, sondern sie brachten mit ihren Urteilen vielmehr ihre Unsicherheit zum Ausdruck (Chapman & Robbins, 1990; Dickinson et al., 1984; Shanks & Dickinson, 1987). Eine analoge Erklärung kann auch für den „Overshadowing“-Effekt entwickelt werden: Auch in diesem Paradigma ist eine Konditionalisierung auf die jeweils andere Ursache nicht möglich (vgl. Waldmann, 2001).

Die Ergebnisse anderer Studien deuten darauf hin, dass die Kontrolle zusätzlicher Ursachefaktoren von gewissen Bedingungen abhängig ist. So wurden weitere Faktoren nur dann berücksichtigt, wenn diese salient und die Anfordernungen an die Verarbeitungskapazität gering waren (Fiedler, Brinkmann, Betsch & Wild, 2000; Schaller & O'Brien, 1992; Spellman et al., 2001; Waldmann & Hagemayer, 2001). Ein weiterer wichtiger Faktor war der kausale Status der zusätzlich vorhandenen Faktoren (Waldmann & Hagemayer, 2001). Probanden konditionalisieren nur auf Faktoren, die nach dem der Aufgabe zu Grunde liegenden Kausalmodell für den Effekt direkt kausal relevant waren. Weitere Befunde hierzu finden sich im Abschnitt über Kausalmodelle.

*Theoretische Kritik an multi-faktoriellen Kontingenztheorien:* Multi-faktorielle Kontingenztheorien liefern unter bestimmten Zusatzannahmen Lösungen für die Probleme der einfachen Kontingenztheorien. So können echte von scheinbaren Kausalrelationen dadurch unterschieden werden, dass echte Kausalrelationen auch bei der Konditionalisierung auf alle denkbaren Kombinationen anderer kausal relevanter Faktoren positive oder negative Kontingenzen mit dem Effekt aufweisen. Im Gegensatz dazu verschwindet die Kontingenz scheinbarer Kausalrelationen durch Konditionalisierung (Ellis, 1991; Waldmann & Hagemayer, 2001). Diese Lösung setzt jedoch voraus, dass auf die richtigen Faktoren konditionalisiert wird. Nehmen wir erneut das Beispiel des Barometers und des Gewitters. Angenommen, wir wollen den Einfluss des Luftdrucks auf die Bildung von Gewittern untersuchen. Da aber der Barometerstand ebenfalls mit dem

Vorhandensein von Gewittern kovariert, berücksichtigen wir diesen als zusätzlichen Faktor. Wir ermitteln also den Zusammenhang von Luftdruck und Gewittern getrennt bei hohem und niedrigem Barometerstand. Das Ergebnis ist, dass wir keinen Zusammenhang zwischen Luftdruck und Gewittern finden werden. Die Erklärung ist, dass es natürlich vollkommen falsch ist, auf den Barometerstand zu konditionalisieren. Eine Konditionalisierung sollte nur auf die für den in Frage stehenden Effekt *kausal relevanten* Alternativenursachen erfolgen (vgl. Cartwright, 1983; Eells, 1991; Waldmann & Hagemayer, 2001). Um kausal relevante von kausal irrelevanten Faktoren unterscheiden zu können, ist es aber notwendig, das Konzept des Kausalmodells einzuführen. Nur innerhalb eines Kausalmodells lässt sich angeben, ob ein Usachefaktor für ein Ereignis kausal relevant ist oder nicht. Wie das obige Beispiel zeigt, reicht eine einfache Kovariation des Faktors mit dem Ereignis als Kriterium nicht aus. Die multifaktoriellen Kontingenzttheorien müssen daher entweder die Begriffe der kausalmodelltheoretischen Ansätze übernehmen (und damit zu Kausalmodelltheorien werden) oder aber sie haben dieselben Probleme wie einfache Kontingenzttheorien.

## 2.2.4 Das probabilistische Kontrast-Modell

Die Theorie der probabilistischen Kontraste (Cheng & Novick, 1990, 1991, 1992; Cheng, Park, Yantis & Holyoak, 1996) entwickelt eine allgemeine, universell einsetzbare, probabilistische Variante des Kovariationsprinzips und verwendet die Kritikpunkte am varianzanalytischen Ansatz aus Gleichzeitigkeit Bestimmung von den einfachen Kontingenzttheorien ab, indem sie annimmt, dass die Bestimmung von Kontingenzen innerhalb einer ausgewählten Teilmenge von Ereignissen vorgenommen wird.

Die Theorie der probabilistischen Kontraste geht davon aus, dass zunächst auf Grund des vorliegenden pragmatischen Kontextes und des inhaltlichen Vorwissens die Menge der Fälle bestimmt wird, innerhalb derer die Ursache-Wirkungszusammenhänge ermittelt werden. Diese Teilmenge der betrachteten Fälle wird als fokaler Set („focal set“) bezeichnet. Mit der Festlegung des fokalen Sets wird entschieden, welche der vorliegenden Bedingungen beim Auftreten eines bestimmten Effektereignisses als konstante Hintergrundbedingungen angesehen werden („causal background“) und welche als mögliche Ursachen hinsichtlich ihrer Kovariation mit dem Effekt näher untersucht werden. Das bedeutet, dass durch die Festlegung des fokalen Sets eine Konditionalisierung auf bestimmte zusätzliche Faktoren vorgenommen wird. So wird beispielsweise für die Untersuchung der Ursachen von Waldbränden angenommen, dass Sauerstoff konstant vorhanden ist. Innerhalb der Fälle, in denen Sauerstoff vorhanden ist, werden dann die Kovariationen von verschiedenen Ursachen wie Dürre oder unerlaubtem Feuermachen und Waldbränden bestimmt.

Die Kovariation der möglichen Ursacheffaktoren mit dem Effekt wird innerhalb des gewählten fokalen Sets in Form von Kontrasten bestimmt. Für die Bestimmung der Auswirkung einzelner Faktoren werden Haupteffektkontraste ermittelt. Diese sind, wie die einfache Kontingenztformel, definiert als die Differenz des Anteils der Fälle, in denen der Effekt nach Auftreten der Ursache auftritt und des Anteils der Fälle, in denen der Effekt auftritt, ohne dass die Ursache zuvor gegeben ist. Je höher der Kontrast ist, desto stärker sollte der geschätzte Kausalinfluss sein. Neben den Haupteffektkontrasten lassen sich auch Interaktionskontraste bestimmen. Diese gehen an, inwieweit der Effekt von der Konjunktion mehrerer Ursachen abhängig ist (siehe Cheng & Novick, 1992, und Cheng, 1993, für Details).

Das Konstrukt des fokalen Sets in Verbindung mit Kontrasten erlaubt es, Ursachen von notwendigen Voraussetzungen zu unterscheiden, den „enabling conditions“. Notwendige Voraussetzungen sind also Bedingungen, die vorliegen müssen, damit die Ursache den Effekt auslösen kann. In dem obigen Beispiel mit dem Waldbrand ist das Vorhandensein von Sauerstoff als notwendige Voraussetzung anzusehen. Nach der Theorie der probabilistischen Kontraste ist die Unterscheidung von Ursache und Voraussetzung eine Folge des gewählten fokalen Sets. Während eine Ursache in einem gegebenen fokalen Set einen Kontrast in Bezug auf den Effekt aufweist, liegt die Voraussetzung in diesem Set konstant als Bedingung vor, kovariert aber in einem anderen fokalen Set mit dem Effekt. Der Sauerstoff in dem Waldbrandbeispiel ist eine notwendige Voraussetzung. Im Gegensatz dazu wird beispielsweise bei einem Brand in einem sauerstofffreien Labor das Vorhandensein von Sauerstoff zu einer Ursache und das Vorhandensein brennbarer Materials zu einer konstanten Bedingung (vgl. auch Hart & Homoté, 1985).

Innerhalb der Theorie ist auch eine Unterscheidung von echten („genuine“) und scheinbaren („spurious“) Ursachen möglich. Eine Ursache ist dann eine scheinbare Ursache für einen Effekt, wenn der in einem fokalen Set vorhandene Kontrast in einem anderen fokalen Set verschwindet, in dem ein anderer kausal relevanter Faktor konstant gegeben ist. So verschwindet die Schein-Korrelation zwischen Barometerstand und Wetter in dem fokalen Set, in dem der atmosphärische Druck konstant gehalten ist.

*Empirische Evidenz:* Die innerhalb eines fokalen Sets bestimmbar Kontraste entsprechen den bedingten Kontingenzen zwischen den jeweiligen Ursacheffaktoren und dem Effektereignis. Daher sind auch alle Befunde, die mit den multifaktoriellen Kontingenzttheorien vereinbar sind, ebenfalls mit der Theorie der probabilistischen Kontraste vereinbar. Bei der Festlegung der fokalen Sets wird das Vorliegen zusätzlicher Ursacheffaktoren berücksichtigt. Je nach Ausprägung dieser Faktoren werden mehrere fokale Sets gebildet und die Kontraste für die einzelnen Sets werden miteinander abgeglichen. Daher sagt diese Theorie auch



den Blockierungseffekt und den Overshadowingeffekt vorher (vgl. Cheng, 1997; Cheng & Novick, 1992; Cheng et al., 1996).

Eine einzigartige, differenzierende Vorhersage liefert die Theorie für die Kausalattribution auf der Grundlage probabilistischer Daten (Cheng & Novick, 1990). Sie sagt vorher, dass Probanden, sofern sie über Informationen bezüglich der Wahrscheinlichkeit eines Effektergebnisses bei allen möglichen Ursachenkombinationen verfügen, ohne Bias eine adäquate Kausalattribution auf Grundlage der vorhandenen probabilistischen Zusammenhänge vornehmen können. In einer breit angelegten Untersuchung (Cheng & Novick, 1990) wurden Probanden daher gebeten, die Ursache für eine Reihe von Ereignissen in der Tradition von Kelley (1967) anzugeben. Die Probanden konnten dabei zwischen dieser Ursachefaktoren (d.h. Interaktionen der Faktoren) wählen. Als Daten erhellend für die Information zu Konsens, Distinktheit und Konsistenz vollständige Informationen über das Auftreten des Effektes bei allen möglichen Kombinationen der Ausprägungen der Ursachefaktoren. Die Befunde bestätigen die Vorhersagen des probabilistischen Konstrastmodells. Wie vorhergesagt, keiner der bekannten Biases auf. Außerdem wählte die Mehrheit der Probanden die auf Grund der vorhandenen Kontraste vorhergesagte Ursache bzw. Ursachenkombination aus. Konkurrierende Theorien kommen als Erklärung ausgeschlossen werden (Cheng & Novick, 1990).

Eine weitere distinkte Vorhersage der Theorie ist die Unterscheidung zwischen Ursache und „enabling condition“, die mit dem Konzept des fokalen Sets erklärt wird. Dazu wurden in einem Experiment (Cheng & Novick, 1991, Experiment 2) Probanden mit vier möglichen Ursachefaktoren (Licht, Wasser, Nährstoffe und Nähe eines Hauses) eines Effektes (Blühen von Dahlien) konfrontiert. Einer der Faktoren war konstant gegeben und kam auf Grund allgemeinen Vorwissens nicht als Ursache in Betracht (Nähe eines Hauses), der zweite Faktor war auf Grund des Vorwissens kausal relevant, war aber ebenfalls konstant, der dritte Faktor wies eine deutlich positive und der vierte eine deutlich negative Kontingenz zum Effekt auf. In verschiedenen Bedingungen wurde die Zuordnung der relevanten Faktoren zu den drei Kontingenzbedingungen manipuliert. Aufgabestimmungen, also anzugeben, ob er Ursache, Inhibitor, notwendige Bedingung oder irrelevant sei. Das Ergebnis war, dass die Probanden jeweils den relevanten, konstant vorhandenen Faktor als notwendige Bedingung ansahen und den relevanten, variierenden Faktor je nach Kontrast als Ursache oder Inhibitor. Der irrelevante Faktor wurde als irrelevant beurteilt.

*Theoretische Kritik der Theorie der probabilistischen Kontraste*: Zwei kritische Punkte teilt die Theorie mit der multifaktoriellen Kontingenztheorie. Der erste

ist, dass nach dieser Theorie ein und dieselbe Ursache sehr viele verschiedene, eventuell engegensetztere Wirkungen auf einen bestimmten Effekt haben kann. Für jedes einzelne fokale Set hat die Ursache einen bestimmten Kontrast, und jeder gibt die Wirkung der Ursache unter den jeweiligen Bedingungen an. Damit gibt es nicht die Wirkung einer Ursache. Dies würde aber bedeuten, dass die Frage nach der Wirkung eines Medikaments streng genommen keinen Sinn macht. Das Medikament hat per se bei jeder Person, ja zu jedem einzelnen Zeitpunkt eine andere Wirkung. Dies widerspricht aber unserer Intuition, dass eine Ursache eine stabile Wirkung auf einen Effekt hat, die bedingungsabhängig von weiteren Faktoren beeinflusst werden kann.

Die zweite kritische Punkt ist, dass es normativ falsch wäre, alle Faktoren, die mit einem bestimmten Ereignis verbunden sind, zu kontrollieren und auf sie zu konditionalisieren (vgl. das obige Beispiel zum Barometerstand und Gewicht). In der Theorie der probabilistischen Kontraste wird davon ausgegangen, dass alle weiteren vorliegenden Faktoren ebenfalls Ursachen des Effektergebnisses sind. Dies muss aber nicht unbedingt der Fall sein.

### 2.2.5 Theorie der Kausalen Power

Die Theorie der kausalen Power („power PC theory“; Cheng, 1997) baut auf der Kritik an früheren kontingenztheoretische Ansätzen auf (siehe oben). Die neue Theorie berücksichtigt nicht nur den Einfluss bekannter und unbekannter Ursachen auf einen Effekt, sondern auch die Intuition, dass eine Ursache eine stabile Wirkung auf einen Effekt hat.

Cheng (1997) geht vom theoretischen Konzept der kausalen Power aus (vgl. auch Cartwright, 1989). Generative kausale Power wird dabei definiert als diejenige Wahrscheinlichkeit, mit der eine Ursache einen Effekt hervorruft, sofern sie vorliegt: „The probability with which c produces e when c is present“ (Cheng, 1997, S. 372). Dabei geht Cheng bei der Konzeption von Power davon aus, dass alle anderen möglichen Ursachen abwesend sind. Inhibitorische Power wird entsprechend als die Wahrscheinlichkeit, mit der eine Ursache zur Abwesenheit eines Effektes führt, sofern sie vorliegt, eingeteilt. An dieser Stelle soll lediglich die theoretische Herleitung für die Schätzung der generativen kausalen Power nachvollzogen werden. Für weitere Details wird auf die Originalarbeit von Cheng (1997) verwiesen. Kausale Power ist ein theoretisches Konstrukt, das sich nicht direkt beobachten lässt, da unter normalen Bedingungen stets weitere unbekanntere Faktoren vorliegen, die den Effekt ebenfalls beeinflussen. Der Idealfall, dass also nur eine Ursache vorliegt, tritt praktisch so gut wie nie ein. Eine Bestimmung der kausalen Power einer Ursache auf der Grundlage von Beobachtungen (oder Handlungen) ist aber unter gewissen Umständen

möglich. Folgende Annahmen werden in der Theorie der kausalen Power gemacht:

1. Eine Ursache C beeinflusst einen Effekt E mit der Power  $P_c$ . Alle weiteren Ursachen, die den Effekt beeinflussen, können zu einem Globalfaktor A zusammengefasst werden und beeinflussen den Effekt mit der Power  $P_a$ .
2. Die Ursachen C und A beeinflussen den Effekt unabhängig voneinander.
3. Die Power beider Ursachen sind unabhängig vom Auftreten der jeweils anderen Ursache.
4. Die Ursachen C und A treten unabhängig voneinander auf.

Die Annahmen 2 und 3 besagen, dass keine Interaktion zwischen den beiden Ursachen vorliegt; die Annahme 4 ist eine Unabhängigkeitsannahme, wie sie in fast allen linearen Modellen gemacht wird (für einen Erweiterung der Theorie auf interaktive kausale Power siehe Novick & Cheng, 2004). Aufbauend auf diese Annahmen ergibt sich die Wahrscheinlichkeit des Effektes der beiden Ursachen C und A als

$$P(c) = P(c) \cdot P_c + P(a) \cdot P_a - P(c) \cdot P_c \cdot P(a) \cdot P_a \quad [1]$$

Die Wahrscheinlichkeit des Effektes resultiert aus der Wahrscheinlichkeit des Auftretens der jeweiligen Ursachen multipliziert mit ihrer kausalen Power vermindert um die Schnittmenge der Fälle, in denen beide Ursachen gemeinsam auftreten. Die Wahrscheinlichkeit des Effektes unter der Bedingung, dass die beobachtete Ursache C aufgetreten ist, lässt sich durch Konditionalisierung auf C bestimmen:

$$P(c|c) = P_c + P(a|c) \cdot P_a - P_c \cdot P(a|c) \cdot P_a \quad [2]$$

Entsprechend ergibt sich für die Wahrscheinlichkeit des Effektes bei Abwesenheit der untersuchten Ursache:

$$P(c|-c) = P(a|-c) \cdot P_a \quad [3]$$

Annahme 4 erlaubt eine weitere Vereinfachung: Wenn die beiden Ursachen C und A unabhängig sind, so gilt:  $P(a|c) = P(a|-c) = P(a)$ . Setzt man diese Vereinfachung in die Gleichungen [2] und [3] ein und berechnet dann die Kontingenz, so ergibt sich Folgendes:

$$\begin{aligned} \Delta P &= P(c|c) - P(c|-c) = P_c + P(a) \cdot P_a - P_c \cdot P(a) \cdot P_a - P(a) \cdot P_a \\ \Delta P &= P_c \cdot (1 - P(a) \cdot P_a) \\ \Delta P &= P_c \cdot (1 - P(c|-c)) \end{aligned}$$

Daraus folgt, dass die Power der untersuchten Ursache C über die beobachtbaren Wahrscheinlichkeiten bestimmt werden kann:

$$P_c = \Delta P / (1 - P(c|-c)) \quad [4]$$

Die obige Formel zeigt, dass die Kontingenz  $\Delta P$  kein adäquater Schätzer des kausalen Einflusses einer Ursache ist. Lediglich, wenn keine weiteren Ursachen den Effekt beeinflussen, d. h. wenn  $P(c|-c) = 0$  ist, stimmen Kontingenz und kausale Power überein. Liegen aber weitere Ursachen vor, lässt sich Formel [4] verwenden, sofern die anderen Ursachen unabhängig auftreten und wirken. Alternativ kann man aber auch nur diejenigen Fälle betrachten, in denen die weiteren Ursachen abwesend sind.

Für inhibitorische Ursachen ergibt eine Ableitung auf Basis derselben Annahmen:

$$P_c = (-1) \cdot \Delta P / P(c|-c) \quad [5]$$

Auch die inhibitorische Power hat, wie die Formel zeigt, nur positive Werte zwischen Null und 1. Ob eine Ursache generativ oder inhibitorisch ist, wird durch das Vorzeichen der Kontingenz bestimmt; inhibitorische Ursachen gehen mit negativen, generative mit positiven Kontingenzen einher.

*Empirische Evidenz.* Eine wesentliche Vorhersage der Theorie der kausalen Power ist, dass Kausalurteile mit zunehmender Wahrscheinlichkeit des Effektes bei gleicher Kontingenz zunehmen. Dies folgt unmittelbar aus der Anwendung von Formel [4]: Nimmt bei konstantem  $\Delta P$  die Wahrscheinlichkeit des Effektes zu (d. h.  $P(c|c)$  und  $P(c|-c)$  steigen an), so wird der Nenner kleiner und damit der Wert der kausalen Power größer. Wie bereits zuvor erwähnt, wurde dieser Effekt bereits vielfach empirisch nachgewiesen (Allan & Jenkins, 1980, 1983; Wasserman et al., 1983; Wasserman, Elek, Charlosh & Baker, 1993). Experimente, in denen sowohl Kontingenzen als auch Powerwerte variiert wurden, unterstützen ebenfalls die Vorhersagen der Theorie der kausalen Power. Der von den Probanden geschätzte Kausaleinfluss einer Ursache nahm mit der beobachtbaren Power zu (Buehner & Cheng, 1997; Buehner, Cheng & Clifford, 2003).

Entgegen der Vorhersage der Theorie der kausalen Power variierten in einigen Experimenten aber auch die Kausalurteile in Abhängigkeit von den präsentierten Kontingenzen bei konstanter kausaler Power (Lohr & Shanks, 2000; White, 2003). Die Unterschiede sind möglicherweise auf die gestellte Aufgabe, insbesondere auf die verwendete abhängige Variable, zurückzuführen. Eine Sensitivität für Kontingenzen resultierte vorwiegend bei Fragen nach dem Zusammenhang von Ursache und Wirkung. Belege für den Einfluss kausaler Power zeigen sich vor allem bei Fragen nach der Auswirkung von möglichen Interven-

tionen oder bei Fragen, die hypothetisch die Bedingung der Abwesenheit der Ko-Faktoren stipulieren (Buchner, Cheng & Clifflord, 2003). Ein weiterer Faktor scheint die Art der Präsentation der Aufgabe zu sein. Formate, die die Informationsverarbeitung erleichtern (z. B. Diagramme), bewirken eher normales Verhalten im Sinne der Power-PC-Theorie als schwierigere Lernaufgaben (z. B. „trial-by-trial“-Lernen).

Eine zweite wesentliche Vorhersage der Theorie der kausalen Power betrifft Randbedingungen, unter denen die generative und inhibitorische Power einer Ursache auf der Grundlage von Daten empirisch ermittelt werden kann. Nach der Formel [4] für den generativen Einfluss einer Ursache ist dieser nicht bestimmbar, wenn die Wahrscheinlichkeit des Effektes auch in Abwesenheit der untersuchten Ursache eins ist ( $P(e|-c) = 1$ ). Ebenso lässt sich die inhibitorische Power einer Ursache gemäß Formel [5] nicht bestimmen, wenn der Effekt auch in Abwesenheit des potenziellen Inhibitors nie auftritt ( $P(e|-c) = 0$ ). Die empirische Gültigkeit dieser Randbedingungen konnte experimentell nachgewiesen werden (Wu & Cheng, 1999). Waren die Randbedingungen nicht erfüllt, so gaben die Probanden an, den Kausaleinfluss der Ursache nicht sicher einschätzen zu können.

Ein weiterer interessanter Testfall für die Theorie der kausalen Power ergibt sich im Zusammenhang mit bedingten Kontrasten. Wie oben erläutert, sagt die Theorie vorher, dass bei Vorliegen einer weiteren generativen Ursache (und keiner weiteren unbekannteren Ursache) die Kontingenz bei Abwesenheit dieser Ursache die kausale Power angibt. Wenn also die auf die Abwesenheit einer alternativen Ursache konditioniertere Kontingenz gleich Null ist, so spricht dies dafür, dass die in Frage stehende Ursache  $C_1$  keinen Einfluss hat,  $\Delta P_c = P(e_1|-c_2) - P(e_1|-c_1) = 0$ . Wenn aber an Stelle dieser bedingten Kontingenz von Ursache gegeben wird und diese gleich Null ist ( $\Delta P_c = P(e_1|-c_2) - P(e_1|-c_1) = 0$ ), so bedeutet dies, dass keine Schätzung des kausalen Einflusses möglich ist. Diese Vorhersage konnte experimentell bestätigt werden (Frattano & Cheng, 1995).

Positive Evidenzen für die Theorie der kausalen Power konnten auch mit anderen lerntheoretischen Paradigmen erbracht werden, etwa für das konditionierte Hemmungsparadigma (Yarlas, Cheng & Holyoak, 1995) und das Overexpectation-Paradigma (Park & Cheng, 1995; vgl. Cheng, 1997, für einen Überblick der Befunde).

*Theoretische Kritik* Die Theorie der kausalen Power ist eine normative, auf den Grundlagen der Wahrscheinlichkeitstheorie basierende Theorie. Wie in der

Herleitung der Theorie dargelegt, muss eine Reihe von Voraussetzungen gemacht werden, die erfüllt sein müssen, damit aus empirischen Daten die kausale Power einer Ursache geschätzt werden kann. Zu diesen zählt die Annahme, dass Ursachen unabhängig voneinander sind und auch unabhängig auf den Effekt wirken. Bis dato wurde nicht überprüft, ob diese Annahmen von Probanden standardmäßig gemacht werden.

Eine Weiterentwicklung der Theorie für Interaktionen zwischen Ursachen wurde vor Kurzem vorgeschlagen (Norrick & Cheng, 2004). Hierdurch wurde die Beschränkung auf unabhängige wirkende Ursachen aufgehoben.

Ebenfalls wurden kürzlich Vorschläge entwickelt, wie die im Rahmen des mechanistischen Ansatzes ermittelten Erkenntnisse über den Einfluss von kausalem Vorwissen durch eine Erweiterung der Theorie erklärt werden können (Lien & Cheng, 2000).

## 2.2.6 Die Theorie der Evidenzevaluation

Die Theorie der Evidenzevaluation (White, 2000, 2002, 2003) sieht im Gegensatz zu anderen regelbasierten Theorien Informationen über das gemeinsame Auftreten von Ursachen und Effekt nicht als Daten, die lediglich die Stärke von Kausalzusammenhängen reflektieren, sondern als Evidenzen, die zur Prüfung einer kausalen Zusammenhangshypothese herangezogen werden können. Beobachtungen, welche mit der Hypothese vereinbar sind, werden als konfirmatorisch, und Beobachtungen, welche nicht mit der Hypothese vereinbar sind, als diskonfirmatorisch angesehen. Es ist ebenfalls möglich, dass gewisse Daten als für die Hypothese irrelevant angesehen werden. Wenn nur eine Ursache und ein Effekt gegeben sind, werden der Theorie zufolge Fälle, in denen die Ursache und der Effekt gemeinsam auftreten oder gemeinsam ausbleiben, als konfirmatorisch gewertet, und Fälle, in denen entweder nur die Ursache oder nur der Effekt auftrifft, als diskonfirmatorisch gewertet. Die Differenz dieser Fälle wird an der Menge aller Beobachtungen normiert. Die resultierende Größe ist der Anteil konfirmatorischer Evidenz („proportion of confirmatory evidence“, pCI, White, 2003). Formal ergibt sich der Anteil konfirmatorischer Evidenz aus den Zellen der Vierfeldertafel (vgl. Abb. 1) als

$$pCI = (A + D - B - C) / (A + D + B + C).$$

Dies ist eine proportionale Version der zuvor vorgestellten AD- oder „Summe der Diagonalen“-Regel. Bei einer Basisrate der Ursache von 0,5 entsprechen die pCI-

Werte den Kontingenzen. Nur bei einer Abweichung von dieser Basisrate werden Unterschiede deutlich.

Der Anteil konfirmatorischer Evidenz bildet der Theorie zufolge die Grundlage für drei Arten von Urteilen: 1. Urteile über das Zutreffen der Hypothese, dass ein Kausalzusammenhang vorliegt; 2. Urteile über die Wahrscheinlichkeit (Likelihood), dass die Ursache den Effekt hervorruft bzw. in einem bestimmten Fall hervorgerufen hat; 3. Urteile über die Stärke des Kausalzusammenhangs. Alle drei Arten von Urteilen sind monoton steigende Funktionen der Anteile konfirmatorischer Evidenz (White, 2003, 2002). Die Theorie lässt sich damit sowohl auf den Erwerb von neuem Wissen über kausale Zusammenhänge als auch auf Kausalattributionen im Einzelfall anwenden. Eine verallgemeinerte Version, die interindividuelle Unterschiede in der Gewichtung der Evidenzen zulässt, wurde ebenfalls entwickelt (White, 2003). Diese Variante der pCI-Regel erlaubt es, auch ideosynkratische Urteilsbildungen zu erfassen. Dadurch wird die Übereinstimmung von Modellierungen der Theorie und empirischen Daten deutlich verbessert (White, 2003).

*Empirische Evidenz.* Wie White (2003) zeigen konnte, ist eine Vielzahl vortiegender experimenteller Befunde (z. B. Kao & Wasserman, 1993; Levin, Wasserman & Kao, 1993; Wasserman, Kao, Van Hamme, Karagiri & Young, 1996) zum Kausallernen mit der Theorie vereinbar. Insbesondere die gefundenen Abweichungen von der Kontingenzregel bei sehr hohen oder niedrigen Basisraten der Ursache können, ebenso wie die häufig gefundene Präferenz von Probanden für die AD-Regel, durch die Theorie erklärt werden.

Des Weiteren liegt eine Reihe von Experimenten mit unterschiedlichen Datenformaten (tabelarisch, in Form von Listen oder „trial-by-trial“) vor. In diesen Studien wurden unterschiedliche Werte für Kontingenz, kausale Power im Sinne von Cheng (1997) und pCI miteinander gekreuzt. Ebenso wurden verschiedene abhängige Variablen untersucht. Die Probanden wurden generell gelehrt, verschiedene Abhängigkeiten der Ursache zu beobachten, wobei die Kausalurteile zunehmend zugegeben, inwieweit die Aussage, dass ein Ereignis ein anderes verursacht, zutrifft. Außerdem wurden sie gelehrt, die Wahrscheinlichkeit einzuschätzen, dass ein Ursache-Wirkungszusammenhang vorliegt, oder einzuschätzen, wie stark der Kausalzusammenhang ist (White, 2003, 2002, 2000). Die Befunde zeigen, dass der Anteil konfirmatorischer Evidenzen all diese Formen von Kausalurteilen beeinflusst. Je höher der Anteil war, desto höher fielen die Urteile aus. Insbesondere konnte gezeigt werden, dass die Kausalurteile zunahm, obwohl die Kontingenz und damit auch die kausale Power konstant bei Null blieben (Experiment 6, White, 2003).

Weitere Belege für die Theorie kommen aus Studien zur Kausalattribution (White, 2002). In diesen Untersuchungen wurden Informationen zu Konsens,

Distinktheit und Konsistenz manipuliert. Die Ergebnisse unterstützen die Vorhersagen der Theorie der Evidenzevaluation im Gegensatz zur Kontingenztheorie.

*Theoretische Kritik.* Die Theorie der Evidenzevaluation ist derzeit eine rein deskriptive Theorie kausalen Denkens und ist als solche in der Lage eine Vielzahl empirischer Befunde zu erfassen und teilweise neue Befunde vorherzusagen. Ein Defizit dieser Theorie ist, dass sie eine Vielzahl freier Parameter aufweist. Jede Art von Evidenz kann nach diesem Modell individuell gewichtet werden. Dies führt dazu, dass eine Vielzahl von Befunden durch entsprechende Gewichtung der Evidenzen erklärt werden kann und damit das Modell schwer zu widerlegen ist. Andere hier vorgestellte Ansätze haben weniger oder gar keine freien Parameter, um eine solche Anpassung vorzunehmen.

Ein weiterer Kritikpunkt ist die geringe Anbindung an normative Modelle. Die Theorie der Evidenzevaluation ist ein hypotheseprüfender Ansatz. Die Hypothese, dass ein bestimmter Kausalzusammenhang vorliegt, wird anhand beobachteter Evidenzen überprüft. Bayesianische Ansätze des Hypothesenprüfens liefern normative Vorgaben, wie eine solche Prüfung erfolgen sollte. Bis dato wurde die Theorie der Evidenzevaluation noch nicht mit diesen normativen Modellen in Verbindung gebracht.

## 2.3 Deduktiver Ansatz

Der deduktive Ansatz betrachtet kausales Denken als eine Form des logischen Denkens. Ausgangspunkt ist die Überlegung, dass Aussagen über Kausalzusammenhänge in der Regel die Form materialer Konditionale haben: Wenn die Ursache vorliegt, dann tritt der Effekt ein. Diese Konditionale repräsentieren im Fall von Kausalität allgemein gültige Kausalgesetze. Nach dem deduktiven Ansatz besteht kausales Denken darin, aus diesem allgemein gültigen Regeln und aus Fakten über das Vorliegen der Ursache bzw. der Wirkung Schlüsse auf das Vorliegen des jeweils anderen Ereignisses zu ziehen.

*Hintergrund.* Der deduktive Ansatz ist eine spezifische Weiterentwicklung von Theorien des konditionalen oder syllogistischen Schließens für den Bereich der Kausalität (vgl. Cummins, 1995; Cummins, Lubart, Alkarnis & Ris, 1991). In der Forschung zum deduktiven Denken hatte sich gezeigt, dass die in den Schlüssen behandelten Inhalte von entscheidender Bedeutung dafür waren, ob die Versuchspersonen einen Schluss als gültig annahmen oder ablehnten (u. a. Cheng & Holyoak, 1985; Evans, 1989; Wason & Johnson-Laird, 1972; vgl. auch Knauff, in diesem Band). So zogen Versuchspersonen bei logischen Schlüssen Hintergrundwissen über den betrachteten Sachverhalt heran und fügten so

impliziert zusätzliche, nicht in der Aufgabenstellung erwähnte Prämissen hinzu. Dies führe dazu, dass das Vorwissen einen höheren Einfluss auf die Entscheidung hatte als die nach der formalen Logik entscheidende syntaktische Form des Schlusses.

Kausalität ist aus mehreren Gründen als ein spezifischer Inhaltsbereich anzusehen. Erstens ist die Kausalrelation inhärent asymmetrisch. Ursachen führen zu ihren Effekten, aber nicht umgekehrt. Dies gilt unabhängig davon, ob die konditionale Aussage mit der Ursache als Bedingung für den Effekt oder mit dem Effekt als Bedingung für die Ursache formuliert wird (Cummins, 1995, 1997). Ein zweiter Grund ist, dass für Schlüsse aus kausalen Prämissen nur bestimmte Arten von Vorwissen relevant sind. So ist das Wissen über alternative Ursachen, generative wie inhibitorische, relevant, das Wissen über zusätzliche Effekte aber nicht. Beispielsweise muss bei einem Schluss vom Vorliegen der Ursache auf das Vorliegen des Effektes die Stärke des Zusammenhangs bzw. das Vorhandensein von möglichen Inhibitoren berücksichtigt werden. Ob eine Ursache hingegen mehrere Effekte verursacht, sollte keinen Einfluss beim Schließen auf einen bestimmten Effekt haben.

*Theorie:* Die deduktive Theorie kausalen Denkens betrachtet Schlüsse aus kausalen Prämissen. Eine der Prämissen ist ein als gültig angenommener Zusammenhang zwischen einer Ursache und einem Effekt in Form einer Wenn-Dann-Aussage. Die zweite Prämisse ist in der Regel eine Aussage über das Vorliegen oder die Abwesenheit der Ursache oder des Effektes. Aus diesen Prämissen soll auf das Vorliegen des jeweils anderen Ereignisses geschlossen werden. Das folgende Beispiel veranschaulicht einen solchen Schluss (nach Cummins, 1995):

Prämisse 1: Wenn die Bremse getreten wird, stoppt das Auto.  
 Prämisse 2: Die Bremse wird getreten.  
 Schluss: Das Auto stoppt.

Im deduktiven Ansatz wird angenommen, dass bei der Ableitung bzw. Überprüfung eines solchen Schlusses inhaltliches Vorwissen über den Sachverhalt in der Prämisse genante Ursache hinreichend und notwendig für den Effekt ist. Eine Ursache ist dann vollständig hinreichend für den Effekt, wenn es keine weiteren Bedingungen oder Faktoren gibt, welche das Eintreten des Effektes trotz Vorliegen der Ursache verhindern. Die Ursache ist unbedingt notwendig für den Effekt, wenn es keine weiteren generativen Ursachen für den Effekt gibt. Basierend auf diesem Vorwissen ergeben sich folgende Implikationen (Cummins, 1995, S. 465):

1. Wenn die Ursache hinreichend für den Effekt ist und die Ursache vorliegt, dann folgt daraus das Vorliegen des Effektes.
2. Wenn die Ursache nicht hinreichend für den Effekt ist und die Ursache vorliegt, dann folgt, daraus nicht das Vorliegen des Effektes.
3. Wenn die Ursache notwendig für den Effekt ist und der Effekt vorliegt, dann folgt daraus das Vorliegen der Ursache.
4. Wenn die Ursache nicht notwendig für den Effekt ist und der Effekt vorliegt, dann folgt daraus nicht das Vorliegen der Ursache.

Diese Implikationen wiederum bestimmen, welcher Schluss aus kausalen Prämissen gezogen und ob ein bestimmter Schluss als gültig angenommen oder abgelehnt wird. Dabei sind die Schlüsse, die auf den genannten Implikationen beruhen, nach Cummins (1995) logisch notwendig. Für kausale Zusammenhänge sind vier Schlusschemata wesentlich: Der Modus Ponens, der Modus Tollens, die Negation des Antecedens und die Affirmation der Konsequenz. Als Kausalzusammenhänge ausgedrückt besagen diese Schlusschemata Folgendes:

*Modus Ponens (MP)*

Prämisse 1: Wenn die Ursache vorliegt, dann tritt der Effekt ein.  
 Prämisse 2: Die Ursache liegt vor.  
 Konklusion: Der Effekt tritt ein.

*Modus Tollens (MT)*

Prämisse 1: Wenn die Ursache vorliegt, dann tritt der Effekt ein.  
 Prämisse 2: Der Effekt ist nicht eingetreten.  
 Konklusion: Die Ursache liegt nicht vor.

*Negation des Antecedens (NA)*

Prämisse 1: Wenn die Ursache vorliegt, dann tritt der Effekt ein.  
 Prämisse 2: Die Ursache liegt nicht vor.  
 Konklusion: Der Effekt tritt nicht ein.

*Affirmation der Konsequenz (AK)*

Prämisse 1: Wenn die Ursache vorliegt, dann tritt der Effekt ein.  
 Prämisse 2: Der Effekt ist eingetreten.  
 Konklusion: Die Ursache liegt vor.

Rein logisch betrachtet sind nur der Modus Ponens und der Modus Tollens gültige Schlüsse; die Negation des Antecedens und die Affirmation der Konsequenz sind logisch nicht zwingend. Der deduktive Ansatz geht aber, wie gesagt, davon aus, dass nicht die logische Gültigkeit des Schlusses für eine Annahme ausschlaggebend ist. Die wahrgenommene Gültigkeit eines kausalen Schlusses wird vielmehr vom Vorliegen anderer kausaler Faktoren bestimmt. Für die Beurteilung des oben vorgestellten Modus Ponens und des Modus Tollens ist demnach relevant, wie hinreichend die Ursache für den Effekt ist: Wenn die Ursache hinreichend ist, muss der Effekt eintreten (MP); außerdem kann eine hinreichende Ursache nicht vorliegen, wenn der Effekt ausbleibt (MT). Die Beurteilung der Negation des Antecedens und der Affirmation der Konsequenz sollte im Gegensatz dazu abhängen, ob die Ursache für den Effekt notwendig ist: Ist die Ursache notwendig, dann folgt aus der Abwesenheit der Ursache die Abwesenheit des Effektes und aus der Anwesenheit des Effektes das Vorliegen der Ursache.

*Empirische Evidenz.* Wie im vorherigen Abschnitt erläutert, sollte die Zustimmung zum Modus Ponens und zum Modus Tollens davon abhängen, inwieweit die Ursache hinreichend für den Effekt ist, während dagegen die Beurteilung der Negation des Antecedens und die Affirmation der Konsequenz von der Notwendigkeit der Ursache bestimmt werden sollte. Ob eine Ursache hinreichend für den Effekt ist, hängt von der Anzahl der Bedingungen ab, die das Eintreten des Effektes verhindern, ob sie notwendig ist von der Anzahl der alternativen generativen Ursachen. Aufbauend auf diese Überlegungen gestaltesten Cummins und Mitarbeiter (1991) ein Experiment, in dem sie die oben vorgestellten Schemata, die Anzahl der bekannten alternativen Ursachen des Effektes und die Anzahl der den Effekt verhöhenden Bedingungen variierten. Letztere Faktoren wurden durch einen Vorwissenstest erhoben und durch eine entsprechende Gestaltung der Aufgabeninhalte manipuliert. Ein Beispiel für eine Ursache, die stark hinreichend, aber nur mäßig notwendig ist, war: „Wenn ich mich in den Finger schneide, blutet er.“ Ein Beispiel für eine weder hinreichende noch notwendige Ursache war: „Wenn ich oft Süßigkeiten esse, bekomme ich Karies.“ Die Probanden bekamen eine Vielzahl verschiedener Schlüsse vorgelegt und sollten angeben, wie sicher sie die jeweilige Schlussfolgerung ziehen können. Das Ergebnis unterstützte die Vorhersagen des deduktiven Ansatzes: Mit steigender Anzahl der den Effekt verhöhrenden Faktoren nahm die Zustimmung zu Modus Ponens und Modus Tollens ab. Die Anzahl alternativer Ursachen hatte keinen kleinen Einfluss auf die Abhebung beim Modus Ponens und nur einen kleinen Einfluss bei Modus Tollens. Engagementsetzte Befunde ergaben sich für die Negation des Antecedens und die Affirmation der Konsequenz. Für diese war, wie erwartet, die Zahl der bekannten weiteren Ursachen des Effektes entscheidend. Dieser Einfluss von Notwendigkeit und Hinänglichkeit auf die Akzeptanz von Schlüssen wurde auch von anderen Forschern, nicht nur für kausal konditionale, nachgewiesen (Thompson, 1994, 2000).

Eine weitere interessante, weil kontraintuitive Vorhersage macht die Theorie für die Auswirkung hoher Vertrautheit des in der Aufgabe dargestellten Sachverhalts: Je besser ein Effekt bekannt ist, desto mehr verschiedene Ursachen können erinnert und desto mehr Fälle, in denen der Effekt trotz Ursache nicht aufgetreten ist, können aus dem Gedächtnis abgerufen werden. Dies führt aber dazu, dass mit hoher Vertrautheit eines Kausalzusammenhangs die Notwendigkeit und Hinänglichkeit der Ursache für den Effekt abnimmt. Dabei sollten kausale Schlüsse bei bekannten Sachverhalten eher abgelehnt, während sie bei unbekanntem Sachverhalten, wie beispielsweise Atomphysik, eher angenommen werden sollten. Diese Vorhersagen konnten experimentell bestätigt werden (Cummins, 1995, Experiment 2).

Bei allen zuvor ausgeführten Schlüssen wurden nur Kausalzusammenhangsprämissen in der Form „Wenn Ursache, dann Effekt“ betrachtet. Es ist aber auch möglich, das Konditional umzustellen, so dass es die Form „Wenn Effekt, dann Ursache“ hat. Logisch betrachtet werden bei der Umstellung Antecedens und Konsequenz vertauscht, kausal gesehen wird zwar die Regel verändert, aber der Kausalzusammenhang belassen. Es hat sich also nichts verändert. Dies zeigen auch die oben hergeleiteten Implikationen. Sie sind unabhängig von der Abfolge von Ursache und Effekt in der ersten Prämisse. Bei diesen wird nur zwischen einem Schluss von der Ursache auf den Effekt und einem Schluss vom Effekt auf die Ursache unterschieden. Daraus lässt sich die Vorhersage ableiten, dass kausal äquivalente Schlüsse gleich beurteilt werden sollten, auch wenn sich ihre syntaktisch-logische Form unterscheidet. Beispielsweise ist der bereits vorgestellte Modus Ponens mit der folgenden Affirmation der Konsequenz kausal, aber nicht logisch gleichwertig (für eine vollständige Liste von kausalen Äquivalenzen siehe Cummins, 1997):

*Affirmation der Konsequenz (2)*  
 Prämisse 1: Wenn der Effekt vorliegt, ist die Ursache aufgetreten.  
 Prämisse 2: Die Ursache ist aufgetreten.  
 Konklusion: Der Effekt liegt vor.

Sowohl beim Modus Ponens als auch bei der obigen Affirmation der Konsequenz muss von der Ursache auf den Effekt geschlossen werden. Damit ist im beiden Fällen relevant, inwieweit die Ursache hinreichend für den Effekt ist. Die Zustimmung zu beiden Schlüssen sollte also von der Anzahl der Bedingungen abhängen, unter denen der Effekt nicht eintritt. Eine umfangreiche experimentelle Untersuchung mit einer Vielzahl unterschiedlicher Schlüsse konnte diese Vorhersagen bestätigen. Nicht die logische Form, sondern die kausale Form

eines Schlusses war für die Urteile der Versuchsteilnehmer ausschlaggebend (Cummins, 1995, Experiment 1).

Untersuchungen anderer Forscher erbrachten nur teilweise bestätigende Daten für den deduktiven Ansatz. So wurde u. a. entgegen den Vorhersagen der Theorie auch ein Einfluss der Anzahl und Stärke alternativer Ursachen auf die Beurteilung des Modus Ponens gefunden (Quinn & Markovits, 1998; Markovits & Porvin, 2001). Die Ergebnisse deuten darauf hin, dass die Stärke der Assoziation zwischen verschiedenen alternativen Ursachen und dem Effekt von Bedeutung ist. Je stärker die anderen Faktoren mit dem Effekt assoziiert sind, desto geringer die Akzeptanz von Modus Ponens und Modus Tollens (vgl. auch Neys, Shaeken & d'Ydewalle, 2002, 2003).

*Kritik:* Der deduktive Ansatz kausalen Denkens bietet eine interessante und eigenständige Sichtweise auf Kausalurteile. Er präzisiert, welche Art von Vorwissen welche Kausalurteile beeinflusst und bringt hierfür eine Reihe überzeugender empirischer Belege. Jedoch stellen die neuen Befunde zum Einfluss der Assoziationsstärke eine Herausforderung dar und fordern dazu auf, den Ansatz zu erweitern.

Weitere kritisch anzumerkende Punkte resultieren vor allem aus der geringen Formalisierung des Ansatzes. Obwohl in der Theorie explizit von Graden der Notwendigkeit und Hinlänglichkeit der Ursache gesprochen wird, bleibt das Konzept vage. Es ist nicht klar, wie die Urteilsbildung erfolgt, wenn eine Ursache weder hinreichend noch notwendig ist. Eine Antwort hierfür könnte die Theorie mentaler Modelle (Johnson-Laird & Byrne, 1991, 2002, vgl. Knauff, in diesem Band) liefern. Nach dieser Theorie wird bei der Aufstellung der mentalen Modelle zu einem Sachverhalt neben den Fällen, in denen Antecedens und Konsequenz gemeinsam vorliegen, berücksichtigt, wie häufig die Konsequenz ohne das Antecedens und das Antecedens ohne die Konsequenz auftritt. So ist es möglich, unterschiedlich notwendige und hinreichende Kausalverbindungen zu repräsentieren, ohne dass numerische Angaben über die Zusammenhänge gegeben sein müssen. Mentale Modelle im Sinne von Johnson-Laird sind trotz ihres qualitativen Charakters hinreichend formalisiert, so dass eine computereinhische Implementierung der Modelle möglich ist (Goldvarg & Johnson-Laird, 2001). Eine andere mögliche Antwort für den Umgang mit nicht hinreichenden und nicht notwendigen Ursachen liefern Bayesianische Ansätze. So haben Sloman und Lagnado (2005) gezeigt, dass sich logische Schlüsse auf der Basis kausaler Prämissen durch diesen Ansatz besser erklären lassen als durch die Theorie der mentalen Modelle (vgl. Abschnitt 2.5 für weitere Details).

Kritisch zu hinterfragen ist auch die deduktive Konzeption des Ansatzes. Die Theorie geht davon aus, dass der Schluss, der auf den Implikationen des Vor-

wissens beruht, deduktiv und damit logisch notwendig ist. Die Konfidenz, welche in diesen Schluss gelegt wird, soll aber von der Notwendigkeit und Hinlänglichkeit der Ursache abhängen und daher unterschiedliche Grade aufweisen. Bisher ist nicht klar, wie solche Grade in diesem Ansatz formalisiert werden sollen. Wie Denise Cummins (1995) selbst zugibt, ist auch eine induktiv-probabilistische Konzeption des kausalen Schließens denkbar. Bei einer solchen Konzeption würde der Schluss auf Grund der Notwendigkeit und Hinlänglichkeit mit einer bestimmten Wahrscheinlichkeit gezogen werden, wie beispielsweise bei Bayesianischen Ansätzen. Eine Entscheidung zwischen beiden Arten von Ansätzen erscheint aber möglich. Der deduktive Ansatz würde vorhersagen, dass sich die Annahme eines Schlusses von der Konfidenz in diesen Schluss trennen lässt, bei induktiven Ansätzen fällt beides zusammen. Entsprechende empirische Befunde würden den deduktiven Ansatz erheblich stärken. Ein weiterer Kritikpunkt an den hier diskutierten Studien ist die Beschränkung auf logisches Denken. Ein verwandter Ansatz, mit dem auch Schlüsse in Lernsituationen modelliert werden können, wurde von De Houwer (De Houwer, 2002; De Houwer & Beckers, 2002a, b; Vandorpe, De Houwer & Beckers, in Druck) entwickelt.

## 2.4 Mechanistischer Ansatz

Der mechanistische Ansatz kausalen Lernens und Denkens betont im Gegensatz zu den assoziativen und regelbasierten Ansätzen die Wichtigkeit inhaltlichen Vorwissens in einer Domäne. Er geht davon aus, dass wissensbasierte Prozesse wesentlich und konstitutiv für kausales Lernen sind.

Dabei greift der mechanistische Ansatz auf eine Vielzahl von Erkenntnissen und theoretischen Modellen aus verschiedenen Bereichen der Psychologie zurück (vgl. Ahn & Bailenson, 1996; Koslowski, 1996). Eine Quelle sind Arbeiten zum Denken und Problemlösen. So zeigte sich, dass beim Denken über physikalische Prozesse häufig qualitative mentale Modelle der physikalischen Mechanismen verwendet werden (Gentner & Stevens, 1986). Fragen über nicht beobachtete Sachverhalte werden häufig durch mentale Simulation auf der Grundlage eines einfachen Modells beantwortet (Hagmayer & Waldmann, 2000; Sloman & Lagnado, 2002, 2005; Wales & Gavanski, 1989). Szenarien und inhaltlich reich ausgestaltete Modelle spielen auch bei Kausalerklärungen im Rahmen von Gerichtsprozessen eine wesentliche Rolle (Pennyngton & Hastie, 1988, 1992).

Weitere Hinweise für die Wichtigkeit kausaler Mechanismen liefert die Forschung zur Repräsentation von Wissensstrukturen. Weltwissen wird häufig in Form von Skripten und Schemata abgespeichert (Schanck & Abelson, 1977; Waldmann, 1990). In diesen Strukturen werden Platzhalter von Ereignissen durch zeitliche, kausale oder funktionale Prozesse verbunden.

Eine weitere Quelle sind Erkenntnisse aus der kognitiven Entwicklungspsychologie (u. a. Bullock et al., 1982; Shultz, 1982; Shultz & Keenan, 1985). Diese sind von allem deshalb interessant, weil das Wissen über kausale Mechanismen mit dem Alter zunimmt. So konnten beispielsweise Bullock und Mitarbeiter (1982) zeigen, dass Kinder bei Kausalurteilen nicht nur auf räumlich-zeitliche Nähe (Konjunktiv) achten, sondern Lücken zwischen Ursachen und Effekt schließen können, wenn ihnen der Mechanismus bekannt ist, der im Verborgenen abläuft.

Ein letzter relevanter Forschungszweig sind Untersuchungen über die sogenannten „folk theories“ (Kozzibir & Keil, 2002; Wilson & Keil, 1998). In diesen Untersuchungen wurde nachgewiesen, dass nicht nur Kinder, sondern auch Erwachsene mit Colledgefahrung beim Lernen und Denken auf intuitive Theorien in verschiedenen Domänen zurückgreifen. Die Theorien sind häufig unvollständig und rudimentär, wobei die Mechanismen im Einzelnen nicht im Detail bekannt sind. Solche skelettierten Theorien finden sich im Bereich der Biologie, der Physik, aber auch der Psychologie. Zusammenfassend gesagt beruht der mechanistische Ansatz auf der Erkenntnis, dass Wissen über kausale Mechanismen, auch wenn dieses eher vage und unvollständig ist, einen großen Einfluss auf Lernen, Denken und Entschieden hat.

*Theorie.* Der mechanistische Ansatz geht davon aus, dass für kausales Lernen und kausale Attribution das Wissen über kausale Mechanismen konstitutiv ist. Weder die Kovariation von Ursache und Wirkung noch zeitliche Priorität der Ursache oder räumlich-zeitliche Konjunktiv sind nach diesem Ansatz hinreichend für einen Kausalzusammenhang. Neben diesen Kriterien muss zusätzlich ein Mechanismus, wenn auch nur rudimentär, kognitiv repräsentiert werden, welcher die Ursache und den Effekt verbindet. Unter Mechanismus verstehen Vertreter des mechanistischen Ansatzes einen Prozess, der dem Zusammenhang zwischen Ursache und Effekt zu Grunde liegt (Ahn et al., 1995, S. 304). Der Mechanismus hat dabei eine „kausale Kraft“ und macht den Effekt „kausal notwendig“ (S. 303). Kausale Mechanismen sind als theoretische Konstrukte anzusehen, die über die beobachtbaren Ereigniszusammenhänge hinausgehen. Das Vorliegen eines Mechanismus macht eine empirisch gegebene Kovariation von Ereignissen zu einer notwendigen Verbindung zwischen Ursache und Wirkung. Damit erklärt das Mechanismuskonzept die weit verbreitete Intuition, dass Kausalzusammenhänge über statistische Relationen hinausgehen und mit Notwendigkeit verbunden sind (vgl. Hume, 1739/1978, 1748/1977; Mackie, 1974).

Eine wesentliche Frage für Vertreter eines mechanistischen Ansatzes ist es, wie spezifisch das Wissen über kausale Mechanismen sein muss, um kausales Lernen und Urteilen zu ermöglichen. Unter Berufung auf die oben angeführten Untersuchungen zu „folk theories“ und rudimentären Kausalmodellen gehen

Vertreter des Ansatzes davon aus, dass bereits grobe Intuitionen über einen möglichen Mechanismus als Grundlage für Kausalattribution und Kausalindikation hinreichend sind. Es ist also ausreichend, wenn auf Grund des gegebenen Vorwissens des Vorliegen eines, wenn auch verborgenen unbekanntes Mechanismus plausibel ist. So sind beispielsweise die elektrischen Prozesse, die der Funktion eines Computers zu Grunde liegen, den meisten von uns unbekannt. Dennoch haben wir die starke Intuition, dass diese vorhanden und notwendig sind. Je spezifischer das Wissen über die kausalen Mechanismen ist, desto eher kann auf der Grundlage von Kovariationsdaten auf das Vorhandensein einer Ursache-Wirkungsbeziehung geschlossen werden.

Auf der Grundlage von Mechanismenwissen kann auch zwischen ethen („genuline“) und scheinbaren („spurious“) Ursachen unterschieden werden. Kovariationen zwischen Ereignissen, für die kein Mechanismus benannt werden kann, sprechen eher gegen das Vorliegen eines Kausalzusammenhangs.

*Empirische Evidenz.* Dass Wissen bzw. Erwartungen über das Vorliegen von Mechanismen für den Erwerb neuen Wissens relevant sind, ist seit Langem bekannt. So konnten beispielsweise Chapman und Chapman (1967, 1969) in einer Reihe von Studien zeigen, dass Annahmen über die Validität von Indikatoren in diagnostischen Verfahren dazu führen, dass in Datensätzen Zusammenhänge der Indikatoren zu klinischen Befunden gesehen werden, die objektiv nicht vorhanden sind. Ähnliche Befunde wurden auch bei praktisch tätigen Therapeuten nachgewiesen (vgl. Fiedler, 2000 für einen Überblick zu Arbeiten zum Einfluss von Vorwissen auf die Wahrnehmung von Zusammenhängen).

Das Vorannahmen die wahrgenommenen kausalen Zusammenhänge spezifisch beeinflussen, konnte in einer Studie zum Einfluss nasser Theorien über die Ursachen von Glück nachgewiesen werden (Kammann & Campbell, 1982). Probanden stellten einen Zusammenhang zwischen plausiblen Faktoren wie Gesundheit und Einkommen und Glück fest, obwohl ein solcher Zusammenhang nicht vorhanden war. Dagegen schätzten sie den Einfluss von Faktoren, zu denen sie keine Vorannahmen hatten, korrekt ein.

In einer Reihe von Studien zum Erwerb von neuem Kausalwissen auf der Grundlage qualitativer Informationen konnte gezeigt werden, dass neben Informationen über das Vorhandensein eines Zusammenhangs und der Größe der Stichprobe, Informationen über mögliche Mechanismen schon bei Kindern von entscheidender Bedeutung sind. Sind Informationen über einen Mechanismus gegeben, so wird die Wahrscheinlichkeit eines Kausalzusammenhangs durchweg höher eingeschätzt (Koslowski, Okagaki, Lorenz & Umbach, 1989; Koslowski, Spilton & Snipper, 1981). Ist das Vorliegen eines Mechanismus plausibel, dann wird bei identischem objektiven Zusammenhang der Kausalfluss signifikant höher eingeschätzt.



In einer der grundlegenden Arbeiten zum mechanistischen Ansatz untersuchen Ahn und Kollegen (Ahn et al., 1995) die Informationssuche bei der Kausalattribution. Während regelbasierte Ansätze davon ausgehen müssten, dass Probanden Informationen über Kovariationen des Ereignisses mit möglichen Ursachen abfragen würden, sagten Vertreter des mechanistischen Ansatzes vorher, dass Probanden ihr Vorwissen aktivieren und überprüfen, welcher der möglichen Mechanismen im untersuchten Fall gegeben war. Um die Informationssuche zu untersuchen, wurden Probanden gebeten, die Fragen zu stellen, die sie zur Ermittlung der Ursache eines Ereignisses beantwortet haben wollen. Die verwendeten Ereignisse knüpften dabei teilweise an die Tradition von Kelley an, gingen aber auch darüber hinaus. Beispielsweise wurden die Probanden damit konfrontiert, dass eine fiktive Mary letztes Jahr in ihrem Job befördert wurde oder dass ein Flugzeug in Nebraska abgestürzt ist. Die von den Probanden gemerkten Fragen wurden danach klassifiziert, ob mit diesen Kovariationen erforscht wurden oder nicht. Das Ergebnis der Studie war, dass sich die Fragen nur zu einem kleinen Teil auf Kovariationen bezogen (zwischen 7 und 13%), mit der überwiegenden Mehrheit der Fragen wurden Hypothesen über bestimmte Mechanismen überprüft (ca. 65%). Probanden wollten beispielsweise lieber wissen, ob Mary besondere Leistungen erbracht hatte, als ob noch andere Personen eine Lohnerhöhung bekommen haben. Der Anteil an Kovariationsfragen erhöhte sich zwar, wenn die Probanden bereits im Vorfeld unvollständige Kovariationsinformationen erhielten (auf ca. 40%), die Kausalattributionen der Probanden bezogen sich aber nach wie vor überwiegend auf Mechanismen (ca. 70%). Mittels Regressionsanalysen konnten die Forscher zeigen, dass sowohl Mechanismus- als auch Kovariationsinformationen einen signifikanten Einfluss auf die Kausalattribution haben. Allerdings erwies sich der Einfluss der Mechanismusinformation als doppelt so stark wie der der Kovariationsinformation.

Weitere Belege für den mechanistischen Ansatz kommen aus Untersuchungen zur Konjunktion von Ursachen bei kausalen Erklärungen eines Ereignisses (Ahn & Bailenson, 1996). Nach regelbasierten und auch kausalmodelltheoretischen Ansätzen steigt die Wahrscheinlichkeit eines Effektes mit der Anzahl unabhängiger Ursachen an. Je mehr Ursachen vorliegen, desto wahrscheinlicher ist der Effekt, sofern keine spezifischen Interaktionen gegeben sind. Nach dem mechanistischen Ansatz wird das Vorliegen eines Ereignisses aber durch einen komplexen, in sich kohärenten Mechanismus erklärt. Ein Wahrscheinlichkeitsanstieg sollte sich daher nur zeigen, wenn sich zusätzliche Ursachefaktoren in den Mechanismus integrieren lassen. Lassen sie sich nicht in den Mechanismus integrieren, so werden sie in ihrer kausalen Relevanz eher abgewertet. Gleiches gilt, wenn die beiden Ursachen Mechanismen implizieren, die nicht gleichzeitig vorliegen können. In diesem Fall sollte ebenfalls einer der Mechanismen abgewertet werden. Die Befunde der Studien deuten darauf hin, dass die Vorträge des mechanistischen Ansatzes zutreffen (Ahn & Bailenson, 1996).

*Theoretische Kritik am mechanistischen Ansatz* Zwei Punkte sind aus theoretischen Gründen kritisch am mechanistischen Ansatz. Der erste Punkt betrifft die Frage, wie Personen zum Wissen über kausale Mechanismen kommen, wenn dieses konstruktiv sprich notwendig für den Erwerb neuen Kausalwissens ist. Cheng (1993) argumentiert beispielsweise, dass Mechanismustheorien erklären, wie einmal erworbenes Wissen über Mechanismen auf neues Lernen übertragen wird, dass sie aber keine plausible Antwort auf die Frage geben, woher das Mechanismusswissen kommt. Cheng (1993) geht davon aus, dass Kovariationen die primäre Quelle für den Erwerb von Neuwissen sind, und dass die Nutzung dieses Wissens auf weiteres Lernen eher nachgeordnet ist. Hierzu wird von Vertretern des mechanistischen Ansatzes häufig auf entwicklungspsychologische Erkenntnisse verwiesen (vgl. Ahn et al., 1995; Koslowski, 1996). Forschungsarbeiten in diesem Bereich deuten darauf hin, dass grundlegende Konzepte und rudimentäres Mechanismusswissen angeboren sein könnten (Baillargeon, 1986, 1987; Gelman, 1990; Leslie & Keeble, 1990; Waldmann, 1996a). Späterer Wissenserwerb über Mechanismen könne auf diesen einfachen Grundstrukturen sukzessive aufbauen.

Der zweite kritische Punkt betrifft die Spezifität oder den Abstraktionsgrad des Mechanismusswissens (vgl. Lien & Cheng, 2000). Um genuine von scheinbaren Ursachen auf der Grundlage von Mechanismussvorwissen unterscheiden zu können, muss dieses Wissen auf einem entsprechenden Abstraktionsgrad vorliegen. So benötigt man Kenntnisse über die zytobiologischen Prozesse im Gastrointestinal-Bereich, um erklären zu können, warum säurehaltige Schmerzmittel und nicht Vitamin C Magenschwüre hervorrufen. Dagegen muss eine Person Wissen über Gravitation haben, um den Kausalzusammenhang zwischen dem Mond und dem Tidenhub richtig als solchen verstehen zu können. Dies ist Wissen auf einem vergleichsweise hohen Abstraktionsgrad. Stimmen Abstraktionsgrad des Vorwissens und der Fragestellung nicht überein, so können nach diesem Ansatz echte von scheinbaren Ursachen nicht unterschieden werden. Wie die Forschung gezeigt hat (vgl. Rozenblit & Keil, 2002) sind auch bei Erwachsenen die meisten Theorien nur rudimentär. Es beschränkt häufig allenfalls eine vage, oft fehlerbehaftete Intuition über das Bestehen eines Mechanismus. Daher müssten Probanden aber fast alle Kovariationen für plausible Kausalzusammenhänge halten, da ihnen das für eine Abgrenzung notwendige Spezialwissen fehlt. Das scheint aber empirisch nicht der Fall zu sein. An dieser Stelle scheint folglich eine genauere Fassung des Mechanismuskonzepts notwendig.

## 2.5 Urteilstheoretischer Ansatz

Die Theorie der kausalen Urteilsbildung von Einhorn und Hogarth (1986) ist eine rein deskriptive Theorie kausalen Denkens in der Tradition Brunswikianscher Modelle der Urteilsbildung (Brunswik, 1952). Kausales Denken bezieht

sich dieser Theorie zufolge auf Merkmale, die sich der direkten Beobachtung entziehen, aber auf Grund von Indikatoren erschlossen werden können. So zeigt sich die kausale Wirkung eines Ereignisses beispielsweise in seiner zeitlichen Priorität und der Kovariation mit dem vermeintlichen Effekt (siehe unten für eine vollständige Darstellung der Indikatoren). Die einzelnen Indikatoren sind dabei fällibel und können im Einzelfall den tatsächlichen Geschehnissen widersprechen. Daher ist kein einzelner Indikator für sich genommen hinreichend. Dennoch kann durch eine Berücksichtigung mehrerer bezüglich ihrer Validität gewichteter Indikatoren auf das Vorliegen einer Ursache-Wirkungsbeziehung zurück geschlossen werden.

Neben diesem allgemeinen Modell der Urteilsbildung baut der Ansatz auch auf philosophischen Überlegungen zur Kausalität, insbesondere dem bedingungsanalytischen Ansatz sowie den psychologischen Spekulationen von Mackie (1974), auf (vgl. 1.1). So geht Einhorn und Hogarth's Theorie davon aus, dass stets eine große Vielzahl von Bedingungen gegeben sein muss, damit ein bestimmtes Ereignis eintritt, von diesen in der Regel aber nur eines als Ursache bezeichnet wird.

Zur Bestimmung der in Frage kommenden Indikatoren für Kausalität werden sowohl philosophische Überlegungen (Hume, 1739/1948; Mill, 1851/1968) als auch Erkenntnisse der Entwicklungspsychologie (Shultz, 1982) und der kognitiven Psychologie (Nisbett & Ross, 1980) herangezogen.

**Theorie:** Die Theorie der kausalen Urteilsbildung bezieht sich auf diagnostische Urteile, das heißt Urteile über die kausale Relevanz möglicher Ursachen für ein zu erklärendes Ereignis. Die Brutto-Stärke („gross strength“) einer solchen Kausalklärung, das heißt die angenommene Stärke des Kausaleinflusses einer möglichen Ursache, wird nach dieser Theorie anhand von sieben Indikatoren bestimmt. Der erste Indikator drückt aus, inwieweit sich der in Frage stehende Faktor von anderen Bedingungen abhebt („degree of difference-in-the-background“  $Q_0$ ). Auf Grund von Kontextfaktoren wie der gewählten Fragestellung, der Perspektive und dem Vorwissen des Urteilers wird ein Teil der Bedingungen den kausal irrelevanten Hintergrundbedingungen zugeordnet und nicht weiter berücksichtigt. Andere Faktoren werden dagegen als mögliche Ursache angesehen. Beispielsweise suchen Pathologen nach anderen Toxikursachen als Mitglieder einer Mordkommission, obwohl beide den Tod kausal erklären wollen. Als Hintergrundfaktoren werden vor allem permanent vorliegende Bedingungen oder übliche Ereignisse angesehen. Diese werden mit Null gewichtet und fallen damit aus der Analyse heraus. Dagegen erhalten ungewöhnliche oder abnormale Ereignisse hohe Gewichtungen. Der zweite Indikator ist die Kovariation zwischen dem betrachteten Faktor und dem beobachteten Effekt ( $Q_1$ ). Eine hohe Kovariation spricht für das Vorliegen eines Kausalzusammenhangs,

ist aber weder notwendig noch hinreichend. Sofern Daten vorhanden sind, kann die Kovariation anhand der beobachteten Fälle der Vier-Felder-Tafel ermittel werden (siehe Abbildung 1). Die einzelnen Fälle können dabei gewichtet werden, je nachdem welcher Aspekt von Kausalität aktuell besonders relevant ist (bspw. Notwendigkeit oder Hinlänglichkeit). Der dritte Indikator für Kausalität ist der Theorie zufolge die zeitliche Abfolge ( $Q_2$ ): Ursachen müssen vor ihren Effekten auftreten. Im Gegensatz zur Kovariation ist dieser Indikator notwendig und kann nicht durch andere Indikatoren kompensiert werden. Der vierte Faktor ist die räum-zeitliche Konjunktivität von Ursache und Effekt ( $Q_3$ ). Eine hohe Konjunktivität spricht für das Vorliegen eines Kausalzusammenhangs. Der fünfte Indikator ist die Ähnlichkeit zwischen Ursache und Effekt ( $Q_4$ ). Beispielsweise verursacht die lineare Rollbewegung einer Billardkugel eher erneut eine lineare Bewegung einer anderen Kugel als eine Rotationsbewegung. Ebenso haben Eltern eher Kinder, die ihnen gleichen, als solche, die ihnen unähnlich sind. Eine große Unähnlichkeit von Ursache und Wirkung spricht daher eher für eine simple Koinzidenz dieser Ereignisse. Der sechste Indikator ist die Kongruenz von Größe und Dauer bei Ursache und Wirkung ( $Q_5$ ). Stimmen diese überein, so spricht dies für einen Kausalzusammenhang. Dagegen erscheint es unwahrscheinlich, wenn beispielsweise eine langsame Bewegung eine sehr schnelle hervorruft, oder ein kleiner Impuls ein großes Gebäude zum Einsturz bringt. Der letzte Indikator ist die angenommene Stärke einer kausalen Kette ( $Q_6$ ), welche den Ursachefaktor und den Effekt verbindet. Diese Stärke kann entweder auf Grund des Vorwissens bestimmt werden oder aber durch Beobachtung der Kovariation der einzelnen Kettenglieder ermittelt werden. Die Stärke der resultierenden Kette ergibt sich dann als Produkt der Stärke der Einzelglieder. Eine Stärke von Null gibt an, dass kein kausaler Mechanismus bekannt ist, welcher den Faktor mit dem Effekt verbinden könnte. Konjunktivität und Kongruenz sind Indikatoren, welche für das Vorliegen einer kausalen Kette sprechen.

Für die Ableitung eines positiven Kausalurteils sind der Theorie zu Folge drei der sieben Indikatoren notwendig und können nicht kompensiert werden. Ein Faktor wird nur dann als Ursache eines Ereignisses angesehen, wenn der Faktor zu den möglichen Ursachen gerechnet wird ( $Q_0 > 0$ ), er zeitlich vor dem Ereignis vorliegt ( $Q_2 > 0$ ) und ein Kausalzusammenhang plausibel ist ( $Q_1 > 0$ ). Außerdem muss entweder eine Kovariation vorliegen oder zumindest eine Ähnlichkeit von Ursache und Effekt gegeben sein. Die Regel zur Integration der Indikatoren zu einem Kausalurteil kann funktional wie folgt dargestellt werden.

$$\text{Gross Strength: } GS(U \rightarrow E) = Q_1 \cdot Q_2 \cdot Q_3 \cdot (Q_4 \cdot Q_5 + \lambda_2 \cdot Q_6)$$

Die Brutto-Stärke einer Kausalklärung ergibt sich aus dem Produkt der Indikatoren zeitliche Priorität, Disjunktheit, Stärke des angenommenen Kausalzusammenhangs und der Summe aus Kongruenz und Ähnlichkeit. Die beiden

Indikatoren Kongruenz und Kongruenz gehen in den Indikator-Stärke des angenommenen Kausalzusammenhangs mit ein. Die beiden Lambda-Parameter  $\lambda_c$  und  $\lambda_s$  repräsentieren die Aufmerksamkeit, die der Kovariation und Ähnlichkeit je nach Aufgabe und Kontext geschenkt werden.

Wichtig ist, dass die Brutto-Stärke einer Erklärung noch nicht das Vorliegen und die Wirkung anderer möglicher Ursachefaktoren für das zu erklärende Ereignis berücksichtigt. Diese Faktoren werden aber für die Bestimmung der Netto-Stärke („net strength“) einer Erklärung mit herangezogen. Die Netto-Stärke ergibt sich durch die Subtraktion der Summe der gewichteten Brutto-Stärken alternativer Kausalerklärungen desselben Ereignisses ( $Z_k \rightarrow E$ ).

$$\text{Netto-Stärke: } NS(U \rightarrow E) = GS(U \rightarrow E) - \sum_k w_k GS(Z_k \rightarrow E)$$

Daraus folgt, dass die Netto-Stärke einer Erklärung mit der Anzahl alternativer möglicher Ursachen und der Stärke ihres Einflusses zurückgehen sollte.

*Empirische Evidenz.* Die empirische Evidenz für die Nutzung von Indikatoren (Cues) bei der Ermittlung von Ursachen ist vielfältig. Wir werden zunächst die einzelnen Indikatoren betrachten und uns dann der Integration zuwenden.

Für die Nutzung von Distinktheit als Indikator sprechen bereits frühe Befunde von Kahneman und Tversky (1982). Sie konnten zeigen, dass Abnormales eher als Ursache angesehen wird als Normales. So wird beispielsweise ein Frontalzusammenstoß zweier Autos dem Fahrer zugeschrieben, der seine Fahrbahn verlassen hat und nicht dem, der einfach geradeaus gefahren ist. Vielfältige Belege für die Distinktheit als wichtiger Faktor für die Kausalattribution haben Hilton und Slugoski (1986) erbracht, die hierfür eine eigene Theorie vorschlagen, das „abnormal conditions focus model“.

Empirische Evidenz für den Einfluss von Kovariation auf das Erlernen von Kausalzusammenhängen wurde bereits in Abschnitt 2.1 ausführlich dargelegt. Sein Einfluss auf Kausalattribution wurde im Rahmen der regelbasierten Ansätze insbesondere dem „Probabilistic Contrast“-Modell von Cheng und Novick (1990, 1992), besprochen (vgl. 2.3). Es kann kein Zweifel bestehen, dass die Kovariation zweier Ereignisse ein wichtiger Indikator für Kausalität ist. Ebenso gibt es aber auch Belege dafür, dass Kovariation nicht hinreichend für Zuschreibung eines ursächlichen Zusammenhangs ist. Das Wissen über einen zu Grunde liegenden Mechanismus scheint zusätzlich von Bedeutung zu sein. So nehmen Personen gelegentlich Kovariationen zwischen Ereignissen wahr, auch wenn keine Kovariation vorliegt (Chapman & Chapman, 1967, 1969). Befunde für die Verzerrung objektiver Kovariationen in Richtung der Vorurteile liegen ebenfalls vor (Koslowski, 1996).

Dass die zeitliche Priorität ein wichtiger Indikator für Kausalität ist, konnte ebenfalls nachgewiesen werden und ist unumstritten. Hierfür sprechen insbesondere entwicklungspsychologische Befunde (Shultz, 1982; Bullock et al., 1982), aber auch Forschungsarbeiten zur Rolle von Zeit bei Wahrscheinlichkeitsurteilen. Kahneman und Tversky (1982) baten beispielsweise Probanden einzuschätzen, mit welcher Wahrscheinlichkeit ein groß gewachsener Vater einen groß gewachsenen Sohn hat und mit welcher Wahrscheinlichkeit ein groß gewachsener Sohn einen großen Vater hat. Obwohl die Wahrscheinlichkeiten gleich sein müssen, schätzten Probanden die erste Wahrscheinlichkeit (Ursache  $\rightarrow$  Wirkung) höher ein als die zweite (Wirkung  $\rightarrow$  Ursache).

Dennoch scheint zeitliche Priorität im Widerspruch zur Behauptung von Einhorn und Hogarth kein notwendiger Faktor zu sein. Forschungsarbeiten zur Theorie der Kausalmodelle, die in Abschnitt 2.6 ausführlich besprochen werden, zeigen, dass Probanden Kausalzusammenhänge auch erlernen, wenn sie zunächst Informationen über das Vorliegen des Effektes und dann über das Vorliegen der Ursachen erhalten (Waldmann & Holyoak, 1992; Waldmann, 1996). Das Wissen über den zu Grunde liegenden Kausalzusammenhang kann also die zeitliche Priorität überschreiben.

Kongruenz wurde als Faktor vielfach untersucht und die Befunde sprechen dafür, dass Kausalurteile höher ausfallen, wenn Ursachen und Effekte raum-zeitlich nahe sind (Mithore, 1946; Shultz, 1982). Forschungsarbeiten mit Erwachsenen zeigen aber auch, dass das Wissen über kausale Mechanismen auch große zeitliche und räumliche Lücken zwischen Ursachen und Effekt überwinden helfen kann (Buehner & May, 2002; Hagnauer & Waldmann, 2002). Beispielsweise würde niemand von uns erwarten, dass eine Kopfschmerztablette unmittelbar nach ihrer Einnahme wirkt, weil wir wissen, dass der Wirkstoff zunächst über den Magen in den Blutkreislauf gelangen und ins Gehirn transportiert werden muss.

Physikalische Ähnlichkeit und Kongruenz als Indikatoren wurden unseres Wissens nach nicht separat untersucht. Die Forschungsarbeiten von Michotte (1946) zur wahrgenommenen Kausalität gehen aber Hinweist darauf, dass diese Faktoren eine Rolle spielen. In diesen Experimenten stieß ein bewegtes Objekt ein anderes an, woraufhin sich dies in Bewegung versetzte. War die resultierende Bewegung des zweiten Objektes der Bewegung des ersten unähnlich, beispielsweise viel schneller, so wurde kein kausaler Zusammenhang wahrgenommen.

Das Vorhandensein einer kausalen Kette oder eines kausalen Mechanismus ist der Letzte der sieben Indikatoren für Kausalität. Im vorherigen Abschnitt zum mechanistischen Ansatz wurden bereits die Belege für den Einfluss von Me-

chiasmusvorwissen ausführlich besprochen (vgl. 2.5). Dieses Wissen scheint einen wesentlichen Einfluss auf den Erwerb von Kausalwissen und auf die Kausalattribution zu haben. Ob dieses aber notwendig ist, darf bezweifelt werden. So wurden viele Experimente zum Kausallernen durchgeführt, in denen die Probanden neue Zusammenhänge erlernen, aber über keinerlei Vorwissen über Mechanismen verfügen. Die Befunde von Roetzheim und Keil (2002) deuten ebenfalls darauf hin, dass Personen zu beinahe allen Sachverhalten eine oberflächliche, knude Theorie entwickeln können, die einen Zusammenhang plausibel erscheinen lässt.

Einhorn und Hogarth (1986) schlagen auch ein Modell zur Integration der einzelnen Indikatoren für Kausalität vor. Dieses beinhaltet, dass zeitliche Priorität, Distinktheit und Mechanismusvorwissen notwendige Voraussetzungen für ein positives Urteil sind. Die oben aufgeführten Forschungsarbeiten weisen aber darauf hin, dass diese starke Annahme wohl nicht aufrechterhalten werden kann. Vielmehr sieht es so aus, dass die einzelnen Indikatoren sich gegenseitig kompensieren können. So kann das Wissen, dass das zuerst beobachtete Ereignis die Folge einer Ursache ist, die Annahme zeitlicher Priorität aufheben. Fehlendes Mechanismusvorwissen kann durch spontane Annahmen ersetzt werden und so Konsistenz mit beobachteten Kovariationen erzielt werden.

*Theoretische Kritik.* Der besondere Verdienst des urteils-theoretischen Ansatzes liegt darin, bereits sehr früh darauf hingewiesen zu haben, dass es eine ganze Reihe von fallhilfen Indikatoren für Kausalität gibt. Wie die nachfolgende empirische Forschung vieler anderen Autoren gezeigt hat, werden die in der Theorie genannten Faktoren tatsächlich als Indikatoren für Kausalzusammenhänge verwendet.

Zwei wichtige Kritikpunkte sind aber dennoch zu nennen. Der eine hat mit der Beschränkung des Ansatzes auf die Erklärung eines Ereignisses zu tun. Die Indikatoren für Kausalität können auch für die Ermittlung komplexer kausaler Strukturen herangezogen werden (vgl. Lagando, Waldmann, Hagemeyer & Stoman, in Druck). Um herauszufinden, wodurch ein Ereignis verursacht wurde, ist es wichtig, alle möglichen Ursachen für dieses Ereignis zu berücksichtigen. Um herauszufinden, ob eine Handlung zum gewünschten Ziel führt, ist es wichtig, alle möglichen Folgen dieser Handlung zu betrachten. Um herauszufinden, wie ein bestimmtes kausales System funktioniert, ist es wichtig, die direkten Kausalbeziehungen von den durch andere Ereignisse vermittelten statistischen Relationen zu trennen. Für all diese Aufgaben können die oben genannten Indikatoren verwendet werden. Wie statistische Modelle zeigen (vgl. Pearl, 2000; Spirtes et al., 1993/2001), ist es ohne diese zusätzlichen Indikatoren praktisch unmöglich, auf Grund von Kovariationen allein kausale Strukturen zu rekonstruieren.

Der zweite Kritikpunkt betrifft die Integrationsfunktion der Indikatoren. Wie oben dargestellt, ist die von Einhorn und Hogarth (1986) vorgeschlagene Integrationsregel nicht mit den empirischen Befunden vereinbar. Die Indikatoren können sich weitaus stärker gegenseitig kompensieren als ursprünglich angenommen. Neue Theorien für die Integration der Indikatoren wurden also gesucht. Beispielsweise nutzen einige Bayesische Theorien die Indikatoren als „constraints“ für den Aufbau von Hypothesen, welche dann anhand der vorhandenen Daten überprüft werden (vgl. Griffiths & Tenenbaum, 2005). Diese Ansätze werden im nächsten Abschnitt ausführlicher behandelt werden.

## 2.6 Kausalmodelle und Bayes-Netze

Die bisher diskutierten psychologischen Theorien haben sich vorwiegend mit Kausalrelationen befasst, bei denen eine oder mehrere Ursachen einen gemeinsamen Effekt beeinflussen. In der Wirklichkeit haben wir es aber in der Regel mit Situationen zu tun, bei denen eine große Zahl von Ursachen und Effekten miteinander in komplexer Weise verbunden sind. Abbildung 2 zeigt ein einfaches Beispiel für ein solches komplexes Kausalmodell über die Entstehung und die Folgen eines grippeartigen Infektes. Die Knoten repräsentieren dabei Ursachen oder Effekte und die Pfeile die direkten Kausaleinflüsse. Viren (A) und Bakterien (B) sind die Ursachen der Infektion (C). C wiederum ist die Ursache für Fieber (D) und Husten (E).

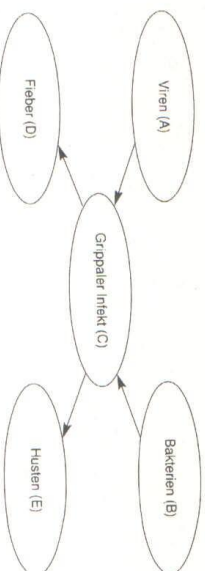


Abbildung 2: Komplexes Kausalmodell der Entstehung und der Folgen grippeartiger Infekte

Eine Reihe von Untersuchungen hat sich in den letzten Jahren mit der Frage befasst, wie wir Wissen über Kausalmodelle erwerben und nutzen. Parallel dazu wurde in der Informatik und Philosophie ein formaler Ansatz zur Repräsentation

nation von Kausalmodellen entwickelt, die so genannten Bayes-Modelle oder Bayes-Netze. Dieser Ansatz wird zunehmend als die normativ richtige Theorie für Kausalität akzeptiert. Eine Reihe von Psychologen und Psychologen hat kausale Bayes-Netze als Werkzeug vorgeschlagen, um menschliches kausales Denken zu modellieren (Glymour, 2001; Gopnik et al., 2004; Rehder, 2003a, b; Seyers, Tenenbaum, Wagenmakers & Blum, 2003; Waldmann & Hagmayer, 2001; Waldmann & Matignon, 1998). Auch wenn sich nicht alle empirischen Studien, die im Folgenden diskutiert werden, selbst als direkte Überprüfung von Bayes-Modellen verstanden haben, lassen sich die Befunde gut anhand dieses Rahmenmodells interpretieren.

**Theorie der kausalen Bayes-Netze.** Kausale Bayes-Netze sind eine Weiterentwicklung der Theorien der probabilistischen Kausalität, deren Defizite sie überwinden (vgl. Abschnitt 1.1) (vgl. Glymour, 2001; Pearl, 1988, 2000; Spirtes et al., 1993/2001). Dieser Ansatz hatte ursprünglich versucht, Kausalität vollständig auf probabilistische Beziehungen zu reduzieren, was zu einer Vielzahl von Problemen geführt hat, da probabilistische Beziehungen im Gegensatz zu kausalen Beziehungen symmetrisch sind. Daher repräsentieren *kausale* Bayes-Netze zunächst die Struktur der kausalen Beziehungen unabhängig von den Daten als kausale Modelle. Ein Beispiel für ein solches Modell ist in Abbildung 2 veranschaulicht. Kausale Bayes-Netze greifen auch grundlegende Ideen aus dem Mechanismus-Ansatz mit auf: Pearl (2000) sieht beispielsweise Kausalmodelle als Repräsentationen von stabilen, miteinander verbundenen kausalen Mechanismen, die allerdings nicht spezifiziert sind (vgl. auch Gopnik et al., 2004). Die kausalen Pfeile sind demnach als Platzhalter für näher zu erforschende Mechanismen anzusehen. Auch Chengs (1997) Power PC-Theorie lässt sich als Spezialfall eines Bayes-Netzes formalisieren, wie Glymour (2001) gezeigt hat.

Die Knoten innerhalb eines Netzwerks stehen für Ursachen oder Effekte und die Pfeile für die kausalen, gerichteten Beziehungen. Das Ereignis, auf das der Pfeil deutet, wird auch häufig Kind, und das am anderen Ende liegende Ereignis Elter genannt. In Abbildung 2 sind etwa A und B die Eltern von C, C ist wiederum der Elter der beiden Kinder D und E. Kausale Bayes-Netze sind azyklische Graphen, das heißt keine Ursache ist vermittelt über andere Ereignisse Ursache ihrer selbst. Dennoch können Wechselwirkungsprozesse durch eine zeitliche Indizierung der Ereignisse als azyklische Graphen repräsentiert werden, so dass die Wirkung einer Ursache die Ursache für einen späteren Zeitpunkt kausal beeinflusst. So führt beispielsweise ein gripplaler Infekt häufig dazu, dass der Erkrankte zu Hause bleibt, wodurch die Wahrscheinlichkeit sinkt, mit Bakterien oder Viren in Kontakt zu kommen. Dies könnte man dadurch repräsentieren, dass man zwei zusätzliche Knoten in obiges Kausalmodell einfügt, die den Kontakt mit Bakterien und Viren zu einem späteren Zeitpunkt repräsentieren, und Pfeile vom gripplalen Infekt auf diese Knoten richtet.

Neben dem Graphen, der die Struktur der Kausalzusammenhänge qualitativ repräsentiert, gehören zu einem vollständigen kausalen Bayes-Netz auch die Parameter des Modells, welche die quantitativen Aspekte erfassen. Zu diesen zählen u. a. die Basistaten der initialen Ereignisse (z. B. die Prävalenz der Viren und Bakterien im obigen Modell) und die Stärke der kausalen Beziehungen (z. B. die Wahrscheinlichkeit mit der ein Infekt zu Fieber führt). Diese Parameter können auf der Grundlage vorhandener Daten geschätzt werden (siehe die Beispiele unten).

Das kausale Strukturmodell hat bestimmte Implikationen für die zu erwartenden statistischen Zusammenhänge zwischen den Ereignissen. So impliziert das in Abbildung 2 gezeigte Modell unter anderem, dass das Auftreten der Viren A mit dem Auftreten von Husten (E) korreliert sein sollte, weil die Viren vermittelt über den gemeinsamen Infekt (C) Husten hervorruft. Abbildung 3A zeigt die implizierten korrelativen als gepunktete Doppelpfeile. Dabei wird deutlich, dass sowohl die Korrelationen als gepunktete Doppelpfeile:  $D, B \rightarrow C \rightarrow E$  miteinander Ereignisse innerhalb einer kausalen Kette (z. B.  $A \rightarrow C \rightarrow D, B \rightarrow C \rightarrow E$ ) miteinander korreliert sind, als auch Effekte einer gemeinsamen Ursache ( $D \leftarrow C \rightarrow E$ ). Dagegen sollten verschiedene Ursachen eines Effektes unkorreliert sein ( $A \rightarrow C \leftarrow B$ ), gegen falls verschiedene Ursachen eines Effektes unkorreliert sein ( $A \rightarrow C \leftarrow B$ ).

Die formale Bedeutung kausaler Bayes-Netze liegt darin, dass sie es ermöglichen, alle aus den Kausalzusammenhängen die Wahrscheinlichkeitsverteilung über alle Ereignisse und damit auch alle zu erwartenden statistischen Zusammenhänge zwischen den Ereignissen zu berechnen. Die zentrale Annahme, die die kausalen Graphen mit Wahrscheinlichkeitsverteilungen verknüpft, ist die Markov-Bedingung. Diese Bedingung besagt, dass jedes Ereignis statistisch unabhängig von allen anderen Ereignissen (außer seinen Kindern) ist, wenn die Werte der Elternvariablen konstant gehalten werden. Anders gesagt, ein Ereignis ist von allen anderen Ereignissen außer seinen Kindern *conditional unabhängig*, wenn seine Elternereignisse bekannt sind. Abbildung 3B veranschaulicht die durch die Markov-Bedingung implizierten bedingten Abhängigkeiten bzw. Unabhängigkeiten am Beispiel des Modells aus Abbildung 2. Die Abhängigkeiten sind als gepunktete Doppelpfeile dargestellt.

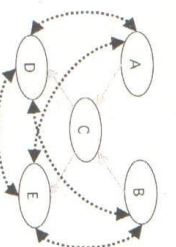


Abbildung 3A. Vom Kausalmodell implizierter statistische Abhängigkeiten.

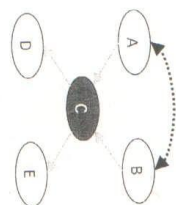


Abbildung 3B. Vom Kausalmodell implizierter konditionale Abhängigkeiten bei Vorliegen von C.

Ein Vergleich der Abbildungen 3A und 3B zeigt die Unterschiede zwischen den implizierten unbefindlichen Abhängigkeiten (3A) und den implizierten bedingten Abhängigkeiten (3B). Kausale Ketten (z.B.  $A \rightarrow C \rightarrow D$ ,  $B \rightarrow C \rightarrow E$ ) implizieren, dass die Ereignisse miteinander korreliert sind, aber bei Vorliegen eines zwischen ihnen liegenden Ereignisses unabhängig voneinander werden: A und D sind unabhängig voneinander, wenn C bekannt ist. Ebenso sind Effekte einer gemeinsamen Ursache korreliert ( $D \leftarrow C \rightarrow E$ ), werden aber unabhängig, wenn die gemeinsame Ursache bekannt ist. Bei verschiedenen Ursachen eines Effektes ist es genau umgekehrt ( $A \rightarrow C \leftarrow B$ ). Diese sollten unbedingte unkorreliert sein, werden aber negativ abhängig, wenn ihr gemeinsamer Effekt gegeben ist. Wenn wir beispielsweise wissen, dass eine Person mit Viren in Kontakt gekommen ist, können wir daraus nicht schließen, ob sie auch mit Bakterien in Kontakt gekommen ist. Wissen wir aber, dass die Person mit Viren in Kontakt war und daraufhin krank geworden ist, so ist es wenig wahrscheinlich, dass sie auch noch mit Bakterien in Kontakt war. Dies ist ein Beispiel für „explaining away“, ein Phänomen, das dem Abwertungsprozess nahe steht, der von Attritionstheoretikern postuliert wird (vgl. Abschnitt 2.2).

Die Markov-Bedingung hat zur Folge, dass die Wahrscheinlichkeitsverteilung über alle Ereignisse innerhalb eines Graphen dem Produkt aller bedingten Wahrscheinlichkeiten, die Kausalzusammenhänge repräsentieren, sowie der Basistraten aller Ursachen, die keine Eltern haben, äquivalent ist:

$$P(\text{Ereignisse}) = \prod_i P(\text{Ereignis}_i | \text{Eltern}_i) \cdot P(\text{Eltern}_i)$$

Das Kausalmodell in Abbildung 2 lässt sich demzufolge folgendermaßen schreiben:

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

An diesem Beispiel kann man die computationale Vereinfachung bereits erkennen, die ein Bayes-Netz ermöglicht. Anstatt die Wahrscheinlichkeiten aller Muster von Ereignissen zu speichern, erlaubt es Bayes-Netze, diese Wahrscheinlichkeiten aus den Parametern des Kausalmodells abzuleiten. So hat das Kausalmodell aus Abbildung 2 mit seinen fünf Ereignissen  $2^5 = 32$  Muster von Ereignissen aber nur 10 Parameter mit deren Hilfe die Wahrscheinlichkeiten aller 32 Muster bei Bedarf errechnet werden können. Diese Vereinfachung kann man sich auch bei Schlüssen zu Nutze machen.

Eine weitere Bedingung, die praktisch das Gegenstück zur Markov-Bedingung ist, ist die Stabilitätsbedingung („stability“, „fäährliness“). Diese Bedingung besagt, dass konditionale Unabhängigkeiten in den Daten durch kausale Graphen und die Markov-Bedingung vorhergesagt werden. Intuitiv besagt diese Bedingung, dass, wenn zwei Ereignisse konditional unabhängig sind, sie durch

keinen direkten Kausalpfad verbunden sind. Praktisch gesehen schließt diese Bedingung beispielsweise Fälle aus, in denen eine Ursache und ein Effekt durch zwei kausale Pfeile verbunden sind, die sich in ihrer Wirkung exakt gegenseitig aufheben, weil der eine Pfeil generativ ist und der andere Pfeil eine in gleicher Stärke inhibitorische Beziehung ausdrückt. Ein solcher Fall von Unabhängigkeit wäre nicht durch die Struktur des Modells, sondern durch die Ausprägung der Parameter bedingt. Diese Stabilitätsbedingung ist wichtig, wenn aus vorliegenden statistischen Zusammenhängen ein Kausalmodell induktiv erschlossen werden soll.

Unterschiedliche Lernalgorithmen wurden für die Rekonstruktion des Strukturmodells aus den beobachteten statistischen Zusammenhängen entwickelt. Eine Variante, die eher top-down vorgeht, sind Bayesische Update-Algorithmen. Liegt kein Vorwissen vor, beginnt man hier mit einer in der Regel uniformen Verteilungshypothese über alle möglichen Graphen (= Kausalmodelle), die man über die Ereignismenge bilden kann. Das heißt, man nimmt zunächst an, dass alle möglichen Kausalmodelle gleich wahrscheinlich sind. Zusätzlich werden auch Vorannahmen über die Parameterverteilung der einzelnen Graphen gemacht. Die Lerndaten werden dann verwendet, um die Plausibilität der Vorannahmen zu revidieren. Dieses Verfahren ist sehr aufwändig, kann aber vereinfacht werden, wenn man über spezifisches Vorwissen verfügt und es mit überschaubaren Domänen tun hat. Waldmann und Marrigton (1998) haben eine Theorie der Kausalmodelle vorgeschlagen, die in diesem Sinne davon ausgeht, dass Versuchsteilnehmer in der Regel Hypothesen über die Struktur und Parametrisierung der Kausalmodelle in die Versuchssituation einbringen, die durch Instruktionen und Vorwissen bedingt sind (vgl. auch Rehder, 2003a, b; Tenenbaum & Griffiths, 2003, in Druck).

Ein eher datengesteuerter Algorithmus sind die so genannten „constraint“-basierten Verfahren. Sie machen sich zu Nutze, dass Bayes-Netze für Tripel von Ereignissen bestimmte statistische Muster vorhersagen (siehe Abbildung 5). So impliziert die Gemeinsamer Effekt-Struktur der Ereignisse A, B und C, dass die Ereignisse A und B unabhängig voneinander auftreten und wirken, aber in Ereignis C statistische Abhängigkeit geraten, wenn die Ausprägung des Effekterignisses C bekannt wird. Betrachtet man dagegen eine kausale Kette, etwa A, C, D, dann impliziert hier die Markov-Bedingung, dass A und D unabhängig voneinander sein sollten, wenn man den Zustand des Elters von D, nämlich C, kennt. Ganz ähnlich sollte es in der Gemeinsame Ursache-Struktur C, D, E aus. Hier sollten D und E unabhängig sein, wenn man C konstant hält. Diese Analyse macht deutlich, dass man Gemeinsamer Effekt-Strukturen auf Grund ihrer spezifischen Implikationen von den beiden anderen Modellen unterscheiden kann. Ketten- und Gemeinsame Ursache-Modelle können aber nicht allein auf Grund der statistischen Zusammenhänge in den Daten unterschieden werden, da sie gleiche

unbedingte und bedingte Abhängigkeiten vorhersagen (Markov-Äquivalenz). Um diese beiden Modelle voneinander zu unterscheiden, benötigt man Zusatzinformationen (etwa über die temporale Abfolge der Ereignisse). Die „constrait“-basierten Algorithmen durchforsten „bottom-up“-Tripel von Ereignissen nach bedingten und unbedingten Abhängigkeiten und konstruieren so Kausalmodelle, die die Daten erklären. Dabei kann es sein, dass mehrere markov-äquivalente Modelle resultieren, die alle gleich gut mit den Daten vereinbar sind.

Eine weitere relativ neue Entwicklung im Bereich kausaler Bayes-Netze ist die Modellierung der Unterschiede von Beobachten und Intervenieren (Pearl, 2000; Spirtes et al., 1993/2001; Woodward, 2003). Diese Analysen gehen von der Beobachtung aus, dass es einen Unterschied macht, ob man das Vorliegen eines Ereignisses lediglich beobachtet oder aktiv durch eine Manipulation dieses Ereignisses gesetzt hat. Sowohl Assoziationstheorien als auch klassische Theorien probabilistischer Bayes-Netze (z. B. Pearl, 1988) machen hier keinen Unterschied. Ein einfaches Beispiel für die Bedeutung dieser Unterscheidung ist der bereits geschilderte Zusammenhang von Barometer und Wetter. Liest man den Wert des Barometers ab, kann man daraus Schlussfolgerungen auf das künftige Wetter ziehen. Versuche ich aber, den gleichen Wert durch Manipulation des Zegers herzustellen, wird dies keinen Einfluss auf das Wetter haben. Ähnlich kann man im Kausalmodell in Abbildung 2 aus der Beobachtung des Effekts D auf den Effekt E schließen. Setzt man hingegen D aktiv, wird E lediglich mit der Wahrscheinlichkeit auftreten, mit der C auftritt und E verursacht. Formalisieren lassen sich Interventionen dadurch, dass man annimmt, dass durch die externe Manipulation alle Einflüsse von den Eltern dieser Variable ausgeschaltet werden. Manipuliere ich D beispielsweise, dann wirkt der Kausalpfad von C auf D nicht mehr, man kann ihn also für die Vorhersage entfernen. Pearl (2000) nennt dies „graph surgery“. Abbildung 4 veranschaulicht die Unterschiede zwischen Beobachtung und Intervention anhand zweier Beispiele auf Grundlage des Kausalmodells aus Abbildung 2.

Da Interventionen die Struktur des Kausalmodells verändern, resultieren häufig unterschiedliche Vorhersagen für Beobachtungen und Interventionen beim selben Kausalmodell. Wie die obige Abbildung zeigt, stimmen die Vorhersagen von Beobachtung und Intervention nur für den Zusammenhang des betroffenen Ereignisses mit ihren (direkten und indirekten) Effekten überein. Dabei ist es völlig irrelevant, wie das Modell und seine Parameter zuvor gelernt wurden. Sind die Parameter aber bekannt, so kann die genaue Wahrscheinlichkeit der anderen Ereignisse bei Beobachtung und Manipulation eines beliebigen Ereignisses innerhalb des Modells berechnet werden.

Interventionen erlauben es auch, zwischen probabilistischen bzw. assoziativen und genuin kausalen Zusammenhängen zu unterscheiden. Kausale Zusammen-

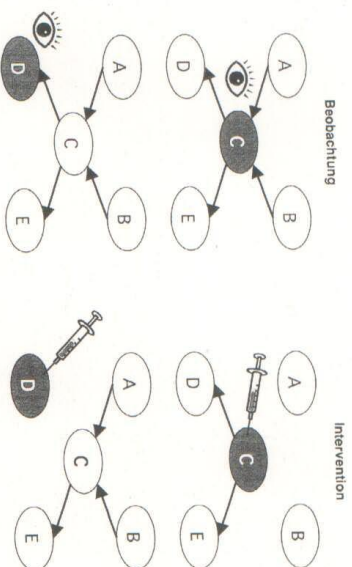


Abbildung 4: Strukturelle Unterschiede von Beobachtungen und Interventionen. Beobachtet bzw. durch Intervention gesetzte Ereignisse sind durch inverteerte Färbung und Augen (Beobachtung) bzw. Spritzen (Intervention) gekennzeichnet.

hänge ermöglichen es, durch eine Intervention ein anderes Ereignis zu verändern, rein statistische Zusammenhänge tun dies nicht. Obwohl gelbe Zähne und Lungenkrebs statistisch miteinander korrelieren, führt das Bleichen der Zähne nicht zu einer Senkung des Lungenkrebsrisikos.

**Empirische Evidenz.** Die ersten Untersuchungen zu Kausalmodellen wurden in dem Bereich Lernen durchgeführt. Waldmann und Holyoak (1992) haben die Frage untersucht, ob sich der Erwerb von Kausalwissen als Assoziation von Ereignissen beschreiben lässt, wie es Assoziationstheorien annehmen, oder ob Versuchsteilnehmer beim Lernen versuchen, strukturierte Kausalmodelle zu lernen. Als Gegenheorie zu Assoziationstheorien wurde eine Theorie der Kausalmodelle vorgeschlagen, die davon ausgeht, dass Versuchsteilnehmer mit hypothetischen Vorannahmen über Kausalmodelle in die Lernsituation gehen, und die Lerndaten dazu verwenden, die Parameter der Modelle zu lernen (z. B. Kausalstärke) (vgl. Waldmann, 1996; Waldmann & Martignon, 1998). Waldmann und Holyoak (1992) und später Waldmann (2000) verwendeten das Blockierungsparradigma, um die beiden Theorien gegeneinander zu testen. Wie im Abschnitt 2.1 und 2.2 bereits geschildert, handelt es sich um eine Lernsituation, in der in Phase 1 ein Cue, etwa eine fiktive Substanz im Blut, gepaart wird mit einem anderen Ereignis, etwa einer fiktiven Blutkrankheit. In Phase 2 wird dann diese Substanz mit einer weiteren Substanz gepaart, so dass beide zusammen die Krank-

heit vorherzusagen. Alle Versuchsteilnehmer lernten die gleiche Aufgabe. Daher erwarten Assoziationsatheorien generell einen Blockierungseffekt: Die Probanden sollten also die kausale Verbindung der zweiten redundanten Substanz anzweifeln.

Die zentrale Manipulation betraf nun den kausalen Status der Lernereignisse. In der prädiagnostischen Lernbedingung wurden die Substanzen als Ursache und die Krankheit als Effekt beschrieben. Hier mussten die Versuchsteilnehmer von den Ursachen prädiagnostiv auf den Effekt schließen lernen. In der diagnostischen Vergleichsbedingung hingegen wurden die gleichen Substanzen als Effekte der Krankheit beschrieben. Hier lernten die Probanden also, die Krankheit auf der Basis ihrer Effekte zu diagnostizieren. Die Befunde zeigten eine interessante Dissoziation. Während es in der prädiagnostischen Bedingung den erwarteten Blockierungseffekt gab, gab es keine signifikante Blockierung in der diagnostischen Bedingung. Dies lässt sich anhand von kausalen Bayes-Netz-Repräsentationen erklären. In der prädiagnostischen Bedingung lernen die Versuchsteilnehmer eine gemeinsame Effekt-Struktur ( $A \rightarrow C \leftarrow B$  in Abb. 2), in der mehrere Ursachen im Wettbewerb um die Erklärung des Effekts stehen. Auf Grund der Kontinuität der beiden Ursachen in Phase 2 lässt sich nichts über die kausale Wirkung der zweiten Ursache aussagen. Anders sieht es in der diagnostischen Bedingung aus. Hier lernen die Probanden eine Gemeinsame Ursache-Struktur ( $D \leftarrow C \rightarrow E$  in Abb. 2). Da die beiden Substanzen Effekte derselben Ursache sind, sollten sie korreliert sein. Das heißt ihr gemeinsames Auftreten in der zweiten Phase war zu erwarten. Da die Wirkung der gemeinsamen Ursache auf jeden Effekt unabhängig von der Wirkung auf den anderen Effekt ist, kann man von jedem Effekt auf die Ursache schließen. Dies gilt unabhängig davon, welche Ausprägung der andere Effekt gerade hat. Beim Lernen der Zusammenhänge zwischen multiplen Effekten und einer gemeinsamen Ursache sollte es daher zu keiner Kompetition kommen.

Die Befunde der Untersuchungen stimmen mit diesen Vorhersagen überein. Weitere bestätigende Befunde für die Theorie der Kausalmodelle wurden auch in einer Overshadowing-Studie von Waldmann (2001) und Untersuchungen von Tangen und Allan (2004) und van Hamme, Kao und Wässeman (1993) gefunden. Es gibt allerdings auch kritische Befunde, die keine Sensitivität für den Aspekt der Kausalrichtung und die Struktur der Kausalmodelle beim Lernen kennen (Cobos et al., 2002; Matur, Arcediario & Miller, 1996; Price & Yates, 1993; Shanks & Lopez, 1996). Derzeit wird vermutet, dass die Komplexität der Aufgabe, die Plausibilität der Instruktionen und die Art der Testung entscheidend dafür sein kann, ob Versuchsteilnehmer in der Lage sind, Wissen über Kausalmodelle adäquat zu erwerben (vgl. Waldmann, 1996, 2000; Waldmann & Holyoak, 1997). Diese Hypothesen wurden in Studien von Waldmann und Walker (2005) untersucht und bestätigt.

Die Entwicklung induktiver, „constrained“-basierter Lernalgorithmen hat einige Forscher inspiriert, der Frage nachzugehen, ob Probanden in der Lage sind, Kausalmodelle auf der Basis von Kovariationen zu induzieren, ohne dass temporale Informationen oder Vorwissen die Klasse plausibler Kausalmodelle einschränkt. Seywers et al. (2003) haben dazu eine Aufgabe entwickelt, bei deren die Versuchsteilnehmer drei fiktive Aliens sahen, die angeblich Gedankenleser waren. Die Probanden wurden in mehreren Lerndurchgängen darüber informiert, welche Gedanken jeder der drei Aliens gerade hatte. Dabei wurde nicht gesagt, wer die Gedanken lesen kann (Effekt) bzw. wessen Gedanken gerade gelesen wurden (Ursache). Die Aufgabe bestand also darin, auf Grund von Gemeinsamkeiten auf das Kausalmodell (Kette: Gemeinsame Ursache oder Gemeinsamer Effekt) zu schließen, das den Mustern zu Grunde lag. Da prinzipiell jeder Alien Ursache oder auch Effekt sein kann, gibt es eine große Vielzahl möglicher Kausalmodelle. Der Befund eines ersten Experimentes war, dass selbst wenn durch die Instruktion nur zwei mögliche Modelle vorgegeben wurden (Gemeinsame Ursache vs. Gemeinsamer Effekt) und die statistischen Muster sehr deutlich für eines der Modelle sprachen, nur eine Minderheit von Versuchsteilnehmern in der Lage war, das Kausalmodell richtig zu erschließen.

Besser wurden die Leistungen, wenn es den Probanden erlaubt wurde zu intervenieren (d. h. Gedanken einzupflanzen). Interessanterweise wählten relativ viele Probanden Interventionen, die es ihnen ermöglichen, zwischen verschiedenen mit den Daten vereinbaren Modellen richtig zu differenzieren. So bevorzugten sie Intervention auf die vermeintlich gemeinsame Ursache bzw. den vermeintlich gemeinsamen Effekt. Wenn dieser Alien die Ursache der Gedanken der anderen war, so sollten diese nach der Intervention derselben Gedanken haben, wenn der Alien aber der Gedankenleser war, so sollte diese Intervention keine Folgen auf die Gedanken der anderen haben. Auch in anderen Studien zeigte sich, dass Interventionen die Menge der möglichen Modelle reduzieren und damit die Induktion des richtigen Modells vereinfachen kann (Gopnik et al., 2004; Lagnado & Sloman, 2002, 2004).

Die Rolle von Kausalmodellen wurde auch im Bereich Kategorisierung untersucht. In einer Reihe von Studien konnte gezeigt werden, dass identische Kategorien-Lernaufgaben zu unterschiedlichem Lernen führen in Abhängigkeit von der eingangs instruierten Kausalstruktur. So haben Waldmann, Holyoak und Fatsane (1995) mit Merkmalen, die kontinuierlich variierten, gezeigt, dass die Versuchsteilnehmer leichter eine nicht-linear trennbare Struktur mit korrelieren Merkmalen lernen, wenn sie der Meinung waren, die Merkmale repräsentieren Effekte; umgekehrt war die linear trennbare Struktur leichter zu lernen, wenn die Lernenden dachten, die Merkmale wären Ursachen. Dieser Befund lässt sich aus der Struktur der kontrastierten Kausalmodelle vorhersagen. Während Gemeinsame Ursache-Modelle eine Korrelation der durch die gemein-



samer Ursache bedingten unabhängigen Effekte implizieren, drücken Gemeinsamer Effekte-Modelle eher die Annahme aus, dass die alternativen Ursachen unabhängig voneinander den Effekt beeinflussen.

Ähnliche Befunde haben auch Rehder und Hastie (2001) mit einem anderen Paradigma gefunden. Sie haben Versuchsteilnehmern Kausalmodelle beigebracht, die unterschiedliche Begriffsmerkmale verknüpfen. So wurde etwa die Kategorie eines Lake Victoria-Strimps mit vier Merkmalen beschrieben, wobei in einer Bedingung (Gemeinsame Ursache-Modell) drei davon (Gewicht, Schlafverhalten, Fluchverhalten) von einem vierten kausal determiniert wurden (ACH-Transmitter). In anderen Bedingungen bzw. Experimenten wurden auch Gemeinsamer Effekte-Modelle und kausale Ketten instruiert (Rehder, 2003a, b; Rehder & Hastie, 2001). Im Anschluss an die Instruktionsphase wurden den Probanden neue Exemplare mit unterschiedlichen Ausprägungen auf diesen Merkmalen vorgelegt mit der Aufgabe einzuschätzen, wie typisch sie für die Kategorien seien. Die Urteile zeigten, dass die Versuchsteilnehmer beeinflusst durch die Kausalmodelle waren. So fanden sie Exemplare tendenziell untypischer, die Merkmalskombinationen aufwiesen, die dem instruierten Kausalszusammenhang widersprachen. Außerdem zeigte sich, dass die Probanden Korrelationen zwischen den Effekten einer gemeinsamen Ursache, aber nicht zwischen mehreren Ursachen eines gemeinsamen Effekts (Gemeinsamer Effekte-Modell) erwarteten (vgl. auch Rehder, 2003a, b). Rehder und Hastie (2001) fanden auch, dass Merkmale ein besonderes Gewicht haben, die mit mehreren anderen Merkmalen kausal verknüpft sind (gemeinsamer Effekt oder gemeinsame Ursache) (vgl. auch Rehder, 2003a, b). Dieser Befund steht im partiiellen Widerspruch zu Ahns (1998) Hypothese, dass Ursachen mehr Gewicht als Effekte haben (vgl. auch Waldmann, 2002, für eine Übersicht über diese Forschung).

Kausalmodelle sind auch wichtig bei der Schätzung kausaler Parameter, z. B. der Stärke der Wirkung einer Ursache. Liegen beispielsweise mehrere alternative, potenziell kontundierte Ursachen vor, ist es angemessen, die potenziellen Störvariablen konstant zu halten, wenn man die Stärke einer Kausalbeziehung abschätzen will. Hingegen sollten in Gemeinsame Ursache-Modellen oder Kausalen Ketten-Modellen weitere Effekte eines Ereignisses nicht konstant gehalten werden, weil dies zu Fehleinschätzungen führt. Waldmann und Hagmayer (2001) haben gezeigt, dass Versuchsteilnehmer diese Richtlinien bei der Einschätzung der Stärke von Kausalbeziehungen tatsächlich intuitiv berücksichtigen (zumindest in nicht allzu komplexen Aufgaben).

Kausalmodelle können komplex sein, so dass häufig nur Wissen über Fragmente erworben wird, die später zusammengeführt werden müssen. Um hier zu korrekten Repräsentationen zu kommen, muss man Wissen darüber haben, wie einzelne Kausalspitze korrekt zu komplexen Kausalmodellen integriert werden

sollten. Hier ist Wissen über die Rolle der Kausalrichtung und die Markov-Bedingung von zentraler Bedeutung. Ahn und Dennis (2000) haben diese Kompetenz für kausale Ketten und Hagmayer und Waldmann (2000), zumindest unter bestimmten Aufgabenbedingungen, für Gemeinsame Ursache- und Gemeinsamer Effekte-Modelle belegt.

Die Unterscheidung zwischen Interventionen und Beobachten wurde in den letzten Jahren ebenfalls von einzelnen Forschergruppen untersucht (siehe Hagmayer, Sloman, Lagrado & Waldmann, in Druck, für eine Übersicht). Die Studien im Bereich des Lernens wurden bereits erwähnt. Sloman und Lagrado (2005) haben kausale Bayesnetze als Modelle logischen Schließens bei kausalen Prämissen vorgeschlagen. In einem Experiment (Experiment 2) haben sie Probanden kausale Prämissen gegeben, die eine kausale Kette beschreiben („Wenn A passiert, dann passiert meistens auch B. Wenn B passiert, dann passiert meistens auch C. A liegt vor. C liegt vor.“). In der Testsituation wurde nun gesagt, es sei beobachtet worden, dass B abwesend sei (Beobachtungsfrage). Alternativ wurde gesagt, dass man aktiv B verhindert habe (Interventionfrage). Die Versuchsteilnehmer unterschieden zwischen diesen Situationen im Sinne kausaler Bayes-Netze. Während sie bei der Beobachtungsfrage auf die Abwesenheit von A schlossen, schlen ihnen die weitere Anwesenheit von A bei der Interventionsfrage plausibel. Da eine Intervention unabhängig von A erfolgt und A nicht beeinflusst, gab es keinen Grund, auf die Abwesenheit des Ereignisses A zu schließen.

Diese Fragestellungen wurden in einer Studie von Waldmann und Hagmayer (2005) weitergeführt und auf den Bereich Lernen übertragen. Die zentrale Frage in den dort berichteten Experimenten ist, ob Versuchspersonen, die Kausalmodelle aus einem mit Daten vermittelten bekommen, aus den Daten die Parameter der Modelle abschätzen und bei Vorhersagen korrekt nutzen können. So erhielten die Probanden beispielsweise die Instruktion, dass sie Wissen über ein Gemeinsame Ursache-Modell erwerben würden (ein Hormon beeinflusst kausal zwei andere Hormone). Im Anschluss an diese Eingangsinstruktion erhielten die Probanden eine Liste von Fällen auf einer Seite, aus der die bedingten Wahrscheinlichkeiten zwischen den Hormonen ermittelbar waren. Die zentrale Manipulation betraf die Testphase: Die Probanden wurden gebeten sich vorzuzubereiten, dass bei 20 Versuchstesten ein Effekte-Hormon beobachtet wurde (Beobachtungsfrage). Die Aufgabe war nun abzuschätzen, wie häufig das andere Effekte-Hormon in dem Gemeinsame Ursache-Modell auftreten würde. Dies ist eine auf Beobachtungen basierende Vorhersage. Die Probanden sollten deshalb von dem beobachteten Effekt zunächst diagnostisch auf die gemeinsame Ursache schließen und dann daraus den zweiten Effekt ableiten, nach dem gefragt wurde. Dies sollte dazu führen, dass die Wahrscheinlichkeit des zweiten Effektes hoch eingeschätzt wird, wenn der erste Effekt anwesend ist, und niedrig-

wenn der erste Effekt abwesend ist. Weitere Fragen bezogen sich auf hypothetische Interventionen. Hier wurden die Probanden gefragt, wie häufig das zweite Effekt-Hormon auftreten würde, wenn man 20 Tieren das erste Effekt-Hormon injizieren würde (Interventionsfrage 1). Ebenso wurden sie danach gefragt, wie häufig das zweite Effekt-Hormon auftreten würde, wenn man 20 Tieren ein Mittel spritzt, das das erste Effekt-Hormon zerstört (Interventionsfrage 2). Da die Probanden in der Lernphase keine Interventionen beobachtet hatten, mussten sie ihre Vorhersage aus den Beobachtungen ableiten. Normativ richtig im Sinne kausaler Bayes-Netzwerke wäre es, den Kausalpfad zwischen der gemeinsamen Ursache und dem durch die Intervention bestimmten ersten Effekt zu streichen („graph surgery“) und die Vorhersage des zweiten Effekts allein darauf zu basieren, wie häufig die gemeinsame Ursache auftritt (Basistrat) und wie stark der zweite Kausalpfad ist. Daher sollten die Probanden die beiden Interventionsfragen gleich beantwortet haben. In einer Reihe von Experimenten konnten Waldmann und Hagemayer (2005) zeigen, dass die Versuchsteilnehmer nicht nur korrekt zwischen Beobachtungs- und Interventionsfragen unterscheiden, sondern dabei auch die Parameter der Modelle (Basistrat; Kausalstärke) korrekt bei der Vorhersage nutzen.

Als letztes Forschungsgebiet soll hier der Bereich induktiven Schließens angesprochen werden. Hier geht es darum, welche Faktoren bestimmen, wie sehr man Informationen über einzelne Individuen auf andere Individuen generalisiert. Ein Beispiel wäre der Schluss von der Prämisse „Alle Korkhähnen haben die Eigenschaft X“ auf die Konklusion „Alle Vögel haben die Eigenschaft X“. Induktive Schlüsse sind nicht zwingend, dennoch scheinen sie uns mehr oder weniger plausibel. Eine Reihe von Modellen wurde entwickelt, die die Ähnlichkeit zwischen den Kategorien der Prämissen und der Konklusion als wichtigen Faktor sehen (vgl. Heit, 2000; Waldmann, 2002, für Überblick, und das Kapitel zum induktiven Schließen von Fiedler & Plesner in diesem Band). In letzter Zeit nehmen aber Ansätze zu, die die kausale Struktur der Kategorien als wichtigen Einflussfaktor sehen.

Rehder und Hastie (2004) haben in einer Studie den Versuchsteilnehmern Kategorien wie die oben schon genannten Lake Victoria-Shrimps beigebracht. Dabei wurden in verschiedenen Bedingungen verschiedene Kausalmodelle (Gemeinsame Ursache-, Gemeinsamer Effekt-, Ketten-Modell) instruiert, die die Merkmale der Kategorien verknüpfen. In der Induktionsaufgabe wurden den Probanden verschiedene Exemplare gezeigt, zusammen mit der Information, dass sie eine bestimmte weitere, neuartige Eigenschaft haben (z. B. „ist von einem Retrovirus infiziert, der das Reproduktionssystem angreift“). Die Probanden sollten daraufhin die Frage beantworten, wie viel Prozent der Exemplare der Kategorie (z. B. der Lake Victoria-Shrimps) ihrer Schätzung nach diese neue Eigenschaft ebenfalls haben würden. Rehder und Hastie konnten anhand von

Kategorien aus einer Reihe verschiedener Domänen zeigen, dass die Versuchsteilnehmer umso stärker generalisieren, je typischer das Exemplar für das Kausalmodell war. Je mehr kausale Beziehungen das Exemplar verletzte, desto geringer waren die angegebenen Prozentwerte. Welches der drei Kausalmodelle verwendet wurde, spielte dabei keine Rolle. Befunde, die in die gleiche Richtung gehen, stammen von Heit und Rubinstein (1994). Sie fanden, dass neuartige biologische Eigenschaften eher von einem Tier auf ein anderes generalisiert werden, wenn man annimmt, dass die Tiere viele biologische Mechanismen teilen.

In den bisher diskutierten Studien zum induktiven Schließen ging es um die Generalisierung neuer Eigenschaften, die bestimmten Individuen zugesprochen wurden. Hadjichristidis, Sloman, Stevenson und Over (2004) haben sich mit der Frage befasst, welche Merkmale von Individuen am ehesten auf andere Individuen übertragen werden. Sie konnten zeigen, dass die Versuchsteilnehmer eher kausale zentrale Eigenschaften generalisieren, von denen viele andere Merkmale abhängen (z. B. gemeinsame Ursachen oder Anfangsglieder von Ketten), als periphere Merkmale. Wurde etwa gesagt, dass viele biologische Funktionen von Eichhörnchen vom Hormon ACTH abhängen, aber nur wenige vom Hormon LH, dann generalisieren die Probanden eher das Hormon ACTH als das Hormon LH auf andere Tiere wie Mäuse. Dieser Effekt war besonders ausgeprägt, wenn die beiden Vergleichskategorien sehr ähnlich waren (z. B. Eichhörnchen – Mäuse), und konnte auch für nicht-biologische Kategorien (Arten) gefunden werden.

Schlieflich haben Medin, Coley, Storms und Hayes (2003) eine Theorie vorgeschlagen, die davon ausgeht, dass induktive Schlüsse auch durch kausale Theorien gestützt werden können. So schließt Probanden von der Prämisse, dass Gras ein bestimmtes Enzym X hat, eher darauf, dass auch Kihle dieses Enzym haben, als darauf, dass auch Menschen dieses Enzym haben, weil man sich leichter vorstellen kann, wie das Enzym vom Gras in die Kihle gelangt. Hier zeigt sich zudem eine kausale Asymmetrie: Den Schluss von Kihlen auf Gras machen die Probanden eher selten, weil der vermutete kausale Mechanismus in die Gegenrichtung läuft. Der Einfluss von bereicherspezifischem Kausalwissen dieser Art wurde auch bei Experten gefunden (Proffitt, Coley & Medin, 2000).

### 3 Schluss

Der Überblick über die Forschung zur Kausalität hat deutlich gemacht, dass es sich um ein höchst produktives Forschungsgebiet handelt. Kausale Kognitionen sind zentral in verschiedenen Domänen und Aufgabenstellungen, so dass ihre Untersuchung eine wichtige Aufgabe für die Kognitionspsychologie ist. Während in der Allgemeinen Psychologie dieses Gebiet lange Zeit vernachlässigt

sigt wurde, zeigt der Einleitungsbeitrag, dass sich die Philosophie schon lange mit diesem Themenbereich befasst hat. Dabei wurde deutlich, dass es eine enge Verbindung zwischen philosophischen und psychologischen Theorien gibt.

Trotz der Vielzahl der Untersuchungsgebiete gibt es auch Forschungslücken. So wissen wir sehr wenig darüber, wie kausales Lernen im Alltag tatsächlich erfolgt. Die meisten Studien untersuchen Lernen in vergleichsweise artifiziellen Situationen. Auch gibt es wenig Untersuchungen dazu, wie wir unser kausales Weltwissen subzessive aufbauen und modifizieren. Ein weiteres interessantes Forschungsgebiet wäre der Einfluss, den Kausalwissen auf Problemlösen hat. In komplexen Problemstellungen geht es häufig um die Vorhersage der Konsequenzen fiktiver Eingriffe in kausale Systeme (vgl. Pearl, 2000, siehe auch Funke, in diesem Band). Eine ebenso neue wie praktisch relevante Frage ist der Zusammenhang von kausalem Denken und Entscheiden. Es erscheint offensichtlich, dass Kausalmodelle bei Entscheidungen herangezogen werden sollten. In der Philosophie wurden auch bereits entsprechende Theorien erarbeitet (Lewis, 1981; Nozick, 1969, 1993), die aber noch ihrer empirischen Überprüfung harren.

Dieses Kapitel hat verdeutlicht, dass derzeit eine Reihe von Theorien miteinander in Konkurrenz steht. Dennoch zeichnet sich eher ein aufeinander Zugeschoben ab als ein Auseinanderstreben. So werden etwa mechanistische und regelbasierte Konzeptionen in die Kausalmodelltheorien erfolgreich integriert. Ob es einmal einen integrativen Ansatz geben wird, der allen Befunden gleichmaßen gerecht wird, wird die Zukunft zeigen. Denkbar sind hier auch Entwicklungen, die von unterschiedlichen aufgabenspezifischen Mechanismen ausgehen, die auf verschiedene Aspekte von Kausalität abheben.

### Literatur

- Ahn, W.-K. (1998). Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition*, 69, 135–178.
- Ahn, W.-K., & Baetens, J. (1996). Causal attribution as a search for underlying mechanisms: An explanation of the conjunction fallacy and the discounting principle. *Cognitive Psychology*, 31, 82–123.
- Ahn, W.-K., & Dennis, M. J. (2000). Induction of causal chains. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Ahn, W.-K., Medin, D. L., & Gelman, S. A. (1993). The role of covariation versus mechanism information in causal attribution. *Cognition*, 54, 299–352.
- Allan, L. G. (1993). Human contingency judgments: Rule-based or associative? *Psychological Bulletin*, 114, 435–448.

- Allan, L. G., & Jenkins, H. M. (1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology*, 34, 1–11.
- Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation*, 14, 381–405.
- Alloy, L. B., & Abramson, L. J. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, 108, 441–485.
- Atkes, H. R., & Hartness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General*, 112, 117–135.
- Audi, R. (1995). *The Cambridge Dictionary of Philosophy*. Cambridge, MA: Cambridge University Press.
- Ballarín, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month old infants. *Cognition*, 23, 21–41.
- Ballarín, R. (1987). Object permanence in 3.5- and 4.5-month old infants. *Developmental Psychology*, 23, 655–664.
- Baker, A. G., Berthier, M., & Vallée-Tourangeau, E. (1989). Judgments of a 2 × 2 contingency table: Sequential processing and the learning curve. *Quarterly Journal of Experimental Psychology*, 41B, 65–97.
- Baker, A. G., Mercier, P., Vallée-Tourangeau, E., 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.
- Berthier, M., & Mercier, P. (1982). Perception of correlation reexamined. *Memory and Cognition*, 10, 511–519.
- Blaisdell, A. P., Bristol, A. S., Gunter, L. M., & Miller, R. R. (1998). Overshadowing and latent inhibition counteract each other: Support for the comparator hypothesis. *Journal of Experimental Psychology: Animal, Behavior Processes*, 24, 335–351.
- Brunswick, E. (1952). *The conceptual framework of psychology*. Chicago, IL: University of Chicago Press.
- Buehner, M. J., & Cheng, P. W. (1997). Causal induction: The power PC theory versus the Rescorla-Wagner theory. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 55–60). Mahwah, NJ: Erlbaum.
- Buehner, M. J., & May, J. (2002). Knowledge mediates the inference of covariation assessment in human causal induction. *Thinking and Reasoning*, 8, 269–295.
- Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1119–1140.
- Bullock, M., Gelman, R., & Ballarín, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York: Academic Press.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.
- Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.

- Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 837–854.
- Chapman, G. B. & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition*, *18*, 537–545.
- Chapman, L. J. & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, *74*, 193–204.
- Chapman, L. J. & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid diagnostic signs. *Journal of Abnormal Psychology*, *74*, 271–280.
- Chudosh, D. L., Neumaner, D. J. & Wasserman, E. A. (1985). Response-outcome contingency: Behavioral and judgmental effects of appetitive and aversive outcomes with college students. *Learning and motivation*, *16*, 1–34.
- Cheng, P. W. (1993). Separating causal laws from causal facts: Pressing the limits of statistical relevance. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 215–264). San Diego, CA: Academic Press.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405.
- Cheng, P. W. & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.
- Cheng, P. W. & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, *58*, 545–567.
- Cheng, P. W. & Novick, L. R. (1991). Causes versus enabling conditions. *Cognition*, *40*, 83–120.
- Cheng, P. W. & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365–382.
- Cheng, P. W., Park, J., Yntas, A. S. & Holyoak, K. J. (1996). A causal-power theory of focal sets. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 36: *Causal learning* (pp. 47–88). San Diego, CA: Academic Press.
- Cobos, P. L., López, F. J., Cano, A., Almaraz, J. & Shanks, D. R. (2002). Mechanisms of predictive and diagnostic causal induction. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*, 331–346.
- Cummings, D. D. (1995). Naïve theories and causal deduction. *Memory & Cognition*, *23*, 646–658.
- Cummings, D. D. (1997). Reply to Earley and Moshkove's comment on 'naïve theories and causal deduction'. *Memory & Cognition*, *25*, 415–416.
- Cummings, D. D., Lubart, T., Aleksis, O. & Riser, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, *19*, 274–282.
- De Houwer, J. (2002). Forward blocking depends on retrospective inferences about the presence of the blocked cue during the elemental phase. *Memory & Cognition*, *30*, 24–33.
- De Houwer, J. & Beckers, T. (2002a). Second-order backward blocking and unovershadowing in human causal learning. *Experimental Psychology*, *49*, 27–33.
- De Houwer, J. & Beckers, T. (2002b). Higher-order retrospective revaluation in human causal learning. *Quarterly Journal of Experimental Psychology*, *55B*, 157–171.

- De Houwer, J., Beckers, T. & Gauthier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, *55A*, 965–985.
- Descartes, R. (1972). *Meditationen über die Grundlagen der Philosophie mit sämtlichen Einwürden und Erwiderungen*. Hamburg: Felix Meiner.
- Descartes, R. (1972). *Meditationen über die Grundlagen der Philosophie mit sämtlichen Einwürden und Erwiderungen*. Hamburg: Felix Meiner.
- Dickinson, A. & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgments. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology*, *49* (B), 60–80.
- Dickinson, A., Shanks, D. R. & Evenden, J. L. (1984). Judgment of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, *36A*, 29–50.
- Domini, M. (2003). *The principles of learning and behavior* (5th ed.). Belmont: Thomson/Wadsworth.
- Dowe, P. (2000). *Physical causation*. Cambridge, UK: Cambridge University Press.
- Edle, E. (1991). *Probabilistic causality*. Cambridge: Cambridge University Press.
- Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*, 3–19.
- Evans, J. S. B. T. (1989). *Bias in human reasoning*. Hillsdale, NJ: Erlbaum.
- Fiedler, K. (1982). Causal schemata: Review and criticism of research on a popular construct. *Journal of Personality and Social Psychology*, *42*, 1001–1013.
- Fiedler, K. (2000). Illusory correlations: A simple associative algorithm provides a convergent account of seemingly divergent paradigms. *Review of General Psychology*, *4*, 25–58.
- Fiedler, K., Brinkmann, B., Betsch, T. & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General*, *129*, 399–418.
- Fincham, F. & Hewstone, M. (2002). Attributionstheorie und -forschung. In W. Stroebe, K. J. Roetzheim, F. & Hewstone (Hrsg.), *Sozialpsychologie* (S. 215–262). Heidelberg: Springer.
- Forsterling, F. (1989). Models of covariation and attribution: How do they relate to the analogy of analysis of variance? *Journal of Personality and Social Psychology*, *57*, 615–625.
- Forsterling, F. (2001). *Attribution: An introduction to theories, research, applications*. Hove, UK: Psychology Press.
- Fratianne, A. & Cheng, P. W. (1995). *Assessing causal relations by dynamic hypothesis testing*. Unpublished manuscript.
- Gelman, R. (1990). First principles organize attention to and learning about relevant data. *Cognitive Science*, *14*, 79–106.
- Gelman, R. & Sevens, A. L. (Hrsg.). (1986). *Mental models*. Hillsdale, NJ: Erlbaum.
- Gilbert, D. T. (1995). Attribution and interpersonal perception. In A. Tesser (Ed.), *Advanced social psychology* (S. 99–147). New York: McGraw-Hill.
- Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: MIT Press.
- Glymour, C. (2003). Learning, prediction and causal Bayes nets. *Trends in Cognitive Sciences*, *7*, 43–48.

- Glymour, C. & Cooper, G. F. (Eds.). (1999). *Computation, causation, and discovery*. Menlo Park, MA: MIT Press.
- Glymour, C. & Johnson-Laird, P. N. (2001). Naive causality: a mental model theory of causal meaning and reasoning. *Cognitive Science*, 25, 565-610.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T. & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 3-32.
- Gopnik, A., Sobel, D. M., Schulz, L. E. & Glymour, C. (2001). Causal learning mechanisms in very young children: Two, three, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, 37, 5, 620-629.
- Griffiths, T. L. & Tenenbaum, J. B. (2005). Elemental causal induction. *Cognitive Psychology*, 51, 334-384.
- Hadjichristidis, C., Sloman, S., Stevenson, R. & Over, D. (2004). Feature centrality and probability induction. *Cognitive Science*, 28, 45-74.
- Hagmayer, Y. & Waldmann, M. R. (2000). Simulating causal models: The way to structural sensitivity. In L. Gleitman & A. Joshi (Eds.), *Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Hagmayer, Y. & Waldmann, M. R. (2002). How temporal assumptions influence causal judgments. *Memory & Cognition*, 30, 1128-1137.
- Hagmayer, Y., Sloman, S. A., Lagnado, D. A. & Waldmann, M. R. (in Druck). Causal reasoning through intervention. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation*. Oxford University Press.
- Hart, H. L. & Honore, A. M. (1985). *Causation and the law*. Oxford: Oxford University Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.
- Heit, E. (2000). Properties of inductive reasoning. *Psychological Bulletin & Review*, 7, 569-592.
- Heit, E. & Rubinsten, J. (1996). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 411-422.
- Hilton, D. J. (1988). Logic and causal attribution. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonense conceptions of causality* (pp. 33-65). Brighton: Harvester Press.
- Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin*, 107, 65-81.
- Hilton, D. J. & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, 93, 75-88.
- Hume, D. (1739/1978). *A treatise of human nature*. Oxford: Clarendon Press.
- Hume, D. (1748/1977). *An enquiry concerning human understanding*. Indianapolis: Hackett Publishing Company.
- Jaspers, J. M. F., Hewstone, M. R. C. & Fincham, F. D. (1983). Attribution theory and research: The state of the art. In J. M. F. Jaspers, F. D. Fincham & M. R. C. Hewstone (Eds.), *Attribution theory: Essays and experiments* (pp. 3-30). Orlando, FL: Academic Press.
- Johnson-Laird, P. N. & Byrne, R. M. J. (1991). *Deduction*. Hove: Lawrence Erlbaum.
- Johnson-Laird, P. N. & Byrne, R. M. J. (2002). Conditionals: a theory of meaning, pragmatics and inference. *Psychological Review*, 109, 646-678.
- Kahneman, D. & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic and A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201-208). New York: Cambridge University Press.
- Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9-33). Miami: University of Miami Press.
- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Handbook and aversive behavior* (pp. 276-296). New York: Appleton-Century-Crofts.
- Kammann, R. & Campbell, K. (1982). Illusory correlation in popular beliefs about the causes of happiness. *Neu Zealand Psychologist*, 11, 52-63.
- Kant, I. (1781). *Kritik der reinen Vernunft*. Frankfurt: Suhrkamp.
- Kao, S.-F. & Wasserman, E. A. (1993). Assessment of an information integration account of confidence: Judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1363-1386.
- Kelley, H. H. (1967). Attribution in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15, S. 192-238). Lincoln, NB: University of Nebraska Press.
- Kelley, H. H. (1972). Causal schemata and the attribution process. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior*. Morristown, NJ: General Learning Press.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, 28, 107-128.
- Kelley, H. H. & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, 31, 457-501.
- Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.
- Koslowski, B., Okagaki, L., Lorenz, C. & Umbach, D. (1989). When covariation is not enough: The role of causal mechanism, sampling method and sample size in causal reasoning. *Child Development*, 60, 1316-1328.
- Koslowski, B., Spilton, D. & Siniper, A. (1981). Children's beliefs about instances of mechanical and electrical causation. *Journal of Applied Developmental Psychology*, 2, 189-210.
- Lachin, H. (1993). *Assoziatives Lernen und Kognition*. Heidelberg: Spektrum Akademischer Verlag.
- Lagnado, D. A. & Sloman, S. A. (2002). Learning causal structure. *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Lagnado, D. A. & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 856-876.
- Lagnado, D. A., Waldmann, M. A., Hagmayer, Y. & Sloman, S. A. Beyond covariation: Cues to causal structure. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation*. Oxford University Press.

- Laplace, P. S. (1814/1912). *Essai philosophique sur les probabilités*. New York, NY: Wiley.
- Larbin, M. J. W., Alden, M. R. F. & Dickinson, A. (1998). Retrospective reevaluation of causal judgments under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1331–1352.
- Leslie, A. M. & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*, 265–288.
- Levin, I. P., Wasserman, E. A. & Rao, S.-E. (1993). Multiple methods for examining biased information use in contingency judgments. *Organizational Behavior and Human Decision Processes*, *55*, 228–250.
- Lewis, D. (1973). *Causation*. Cambridge, MA: Harvard University Press.
- Lewis, D. (1981). Causal decision theory. *Australian Journal of Philosophy*, *59*, 5–30.
- Lewis, D. (1986). *Philosophical Papers*, Vol. II. New York: Oxford University Press.
- Lien, Y. & Cheng, P. W. (2000). Distinguishing genuine from spurious causes: A coherence hypothesis. *Cognitive Psychology*, *40*, 87–137.
- Lohr, K. & Shanks, D. R. (2000). Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, *107*, 195–212.
- Locke, J. (1700). *An essay concerning human understanding*. London: Four Books.
- Lopez, E. J., Almaraz, J., Fernandez, P. & Shanks, D. R. (1999). Adquisición progresiva del conocimiento sobre relaciones predictivas: Curvas de aprendizaje en juicios de contingencia [Gradual acquisition of knowledge of predictive relationships: Learning curves in contingency judgments]. *Psicología*, *11*, 337–349.
- Machio, S. & Burkart, J. (2002). Recursive retrospective reevaluation of causal judgments. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *28*, 1171–1186.
- Mackie, J. L. (1974). *The cement of the universe*. Oxford: Clarendon.
- Markovits, H. & Porvin, F. (2001). Suppression of valid inferences and knowledge structures: The curious effect of producing alternative antecedents on reasoning with causal conditionals. *Memory & Cognition*, *29*, 736–744.
- Marzke, H., Acechano, F. & Miller, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 182–196.
- McArthur, L. Z. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, *22*, 171–193.
- McArthur, L. Z. (1976). The lesser influence of consensus than distinctiveness information on causal attribution. *Journal of Personality and Social Psychology*, *33*, 733–742.
- Medin, D. L., Coley, J. D., Storms, G. & Hayes, B. K. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review*, *10*, 517–532.
- Milcharek, A. E. (1946). *The perception of causality*. New York, NY: Basic Books.
- Mill, J. S. (1851/1968). *System der deduktiven und induktiven Logik*. Aalen: Scientia.
- Miller, R. R. & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 51–92). New York: Academic Press.
- Neuhaber, D. J. & Wasserman, E. A. (1986). The effects of unidirectional versus bidirectional rating procedures on college students' judgments of response-outcome contingency. *Learning and Motivation*, *17*, 162–179.
- Nors, W., Shalcken, W. & d'Ydewalle, G. (2002). Causal conditional reasoning and semantic memory retrieval: a test of the semantic memory framework. *Memory & Cognition*, *30*, 908–920.
- Nors, W., Shalcken, W. & d'Ydewalle, G. (2003). Causal conditional reasoning and strength of association: the disjunctive condition case. *European Journal of Cognitive Psychology*, *15*, 161–176.
- Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.
- Nisbett, R. E., Peng, K., Chao, I. & Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review*, *108*, 291–310.
- Novick, L. R. & Cheng, P. W. (2004). Assessing interactive causal power. *Psychological Review*, *111*, 455–485.
- Nozick, R. (1969). Newcomb's problem and two principles of choice. In N. Rescher (Ed.), *Essays in honor of Carl G. Hempel* (pp. 114–146). Dordrecht: D. Reidel.
- Nozick, R. (1993). *The nature of rationality*. Princeton: Princeton University Press.
- Ovris, B. R., Cunningham, J. D. & Kelley, H. H. (1975). A closer examination of causal inference: The role of consensus, distinctiveness, and consistency information. *Journal of Personality and Social Psychology*, *32*, 605–616.
- Park, J.-K. & Cheng, P. W. (1995). Boundary conditions on overexpectation in causal learning with discrete trials: A test of the power PC theory. Unpublished manuscript.
- Peirce, J. M. (1994). Similarity and discrimination – A selective review and a connectionist model. *Psychological Review*, *101*, 587–607.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufmann Publishers.
- Pearl, J. (2000). *Causation: Models, reasoning and inference*. Cambridge, MA: Cambridge University Press.
- Pennington, N. & Hastie, R. (1988). Explanation based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 521–533.
- Pennington, N. & Hastie, R. (1992). Explaining the evidence: Test of the story model for juror decision making. *Journal of Personality and Social Psychology*, *62*, 189–206.
- Pice, P. C. & Yates, J. F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition*, *21*, 561–572.
- Pice, P. C. & Yates, J. F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1639–1655.
- Poffitt, J. B., Coley, J. D. & Medin, D. L. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 811–838.

- Quinn, S. & Markovits, H. (1998). Conditional reasoning, causality, and the structure of semantic memory: strength of association as a predictive factor for content effects. *Cognition*, 68, B93-B101.
- Rohrer, B. (2003a). Categorization as causal reasoning. *Cognitive Science*, 27, 709-748.
- Rohrer, B. (2003b). A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1141-1159.
- Rohrer, B. & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, 130, 323-360.
- Rohrer, B. & Hastie, R. (2004). Category coherence and category-based property induction. *Cognition*, 91, 113-153.
- Rescorla, R. A. (1968). Probability of check in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, 66, 1-5.
- Rescorla, R. A. (1991). Associations of multiple outcomes with an instrumental response. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 465-474.
- Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- Rozzblat, L. & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of epistamory depth. *Cognitive Science*, 26, 521-562.
- Sahnon, W. C. (1980). Probabilistic causality. *Pacific Philosophical Quarterly*, 61, 50-74.
- Salmou, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.
- Schaller, M. & O'Brian, M. (1992). "Intuitive analysis of covariance" and group stereotype formation. *Personality and Social Psychology Bulletin*, 18, 776-785.
- Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.
- Schusterck, M. W. & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, 110, 101-120.
- Shallice, H. (1983). Human covariation judgment: Accuracy and strategy. *Learning and Motivation*, 14, 433-448.
- Shallice, H. & Mims, M. (1982). Sources of error in judging event covariation: Effects of memory demands. *Journal of Experimental Psychology: Human Learning and Memory*, 8, 208-224.
- Shallice, H. & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory & Cognition*, 8, 459-467.
- Shanks, D. R. (1985). Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology*, 37B, 1-21.
- Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation*, 18, 147-166.

- Shanks, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 433-445.
- Shanks, D. R. (1993). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology*, 84, 519-534.
- Shanks, D. R. & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 21, pp. 229-261). New York: Academic Press.
- Shanks, D. R. & Lopez, F. J. (1996). Causal order does not affect cue selection in human associative learning. *Memory and Cognition*, 24, 511-522.
- Shanks, D. R., Charles, D., Darby, R. J. & Azmi, A. (1998). Configural processes in human associative learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1353-1378.
- Shanks, D. R., Pearson, S. M. & Dickinson, A. (1989). Temporal contiguity and the judgment of causality. *Quarterly Journal of Experimental Psychology*, 41B, 139-159.
- Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, 47 (No. 1).
- Shultz, T. R. & Kesnerbaum, N. R. (1985). Causal reasoning in children. In G. Whitehurst (Ed.), *Annals of Child Development* (Vol. 2, pp. 195-249). Greenwich, CT: JAI Press.
- Stoman, S. A. & Laguarda, D. A. (2002). Counterfactual undoing in deterministic causal reasoning. *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, Mahwah: Erlbaum.
- Stoman, S. A. & Laguarda, D. A. (2005). Do we do? *Cognitive Science*, 29, 3-39.
- Spellman, B. A. (1996a). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science*, 7, 337-342.
- Spellman, B. A. (1996b). Conditionalizing causality. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 34. *Causal learning* (pp. 167-200). San Diego: Academic Press.
- Spellman, B. A., Price, C. M. & Logan, J. M. (2001). How two causes are different from one: The use of (un)conditional information in Simpson's paradox. *Memory & Cognition*, 29, 193-208.
- Spetcher, D., Premack, D. & Premack, A. (1995). *Causal cognition: A multidisciplinary debate*. New York, NY: Clarendon.
- Sprites, P., Glymour, C. & Scheines, R. (1993/2001). *Causation, prediction and search*. New York: Springer.
- Stevvers, M., Tenenbaum, J. B., Wagenmakers, E.-J. & Blum, B. (2003). Inferring causal relations from observations and interventions. *Cognitive Science*, 27, 453-489.
- Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North Holland.
- Suppes, P. (1984). *Probabilistic metaphysics*. Oxford: Basil Blackwell.
- Tangen, J. M. & Allan, L. G. (2004). Cue-interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition*, 32, 107-124.

- Tenenbaum, J. B. & Griffiths, T. L. (2003). Theory-based causal inference. *Advances in Neural Information Processing Systems*, 15, 35-42.
- Tenenbaum, J. B. & Griffiths, T. L. (in press). Intuitive theories and rational causal inference. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation*. Oxford University Press.
- Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22, 742-758.
- Thompson, V. A. (2000). The task-specific nature of domain-general reasoning. *Cognition*, 76, 209-268.
- Vandorpe, S., De Houwer, J. & Beckers, T. (in press). Further evidence for the role of inferential reasoning in forward blocking. *Memory and Cognition*.
- Van Hamme, L. J. & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, 25, 127-151.
- Van Hamme, L. J., Kao, S. F. & Wasserman, E. A. (1993). Judging interevent relations: From cause to effect and from effect to cause. *Memory and Cognition*, 21, 802-808.
- von Wright, G. (1971). *Explanation and understanding*. Ithaca, NY: Cornell University Press.
- Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Hypnotism: processing in animals: Memory mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.
- Wagner, A. R. & Brandon, S. E. (2001). A componential theory of Pavlovian conditioning. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 23-64). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Waldmann, M. R. (1990). *Schemata und Gedächtnis: Das Zusammenwirken von Raum- und Ereigniswissen beim Gedächtnis für Alltagsituationen*. Heidelberg: Asanger Verlag.
- Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 34. *Causal learning* (pp. 47-88). San Diego: Academic Press.
- Waldmann, M. R. (1996a). Wissenschaftleres Lernen. In J. Hoffmann & W. Kinosh (Hrsg.), *Enzyklopädie - Kognition, Band II: Lernen* (S. 323-355). Göttingen: Hogrefe Verlag.
- Waldmann, M. R. (1997). Wissen und Lernen. *Psychologische Rundschau*, 48, 84-100.
- 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. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychological Bulletin*, 127, 600-608.
- Waldmann, M. R. (2002). Kategorisierung und Wissenserwerb. In J. Mühlbacher & W. Prinz (Hrsg.), *Lehrbuch Allgemeiner Psychologie* (S. 432-491). Heidelberg: Spektrum Verlag.
- Waldmann, M. R. & Hagmayer, Y. (2001). Estimating causal strength: The role of structural knowledge and processing effort. *Cognition*, 82, 27-58.
- Waldmann, M. R. & Hagmayer, Y. (2005). Seeing vs. doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 216-227.

- 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. (1997). Determining whether causal order affects cue selection in human contingency learning: Comments on Shanks and Lopez (1996). *Memory & Cognition*, 25, 125-134.
- Waldmann, M. R. & Marrigron, L. (1998). A Bayesian network of causal learning. In M. A. Gershbach & S. J. Dery, *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1102-1107). Mahwah, NJ: Erlbaum.
- Waldmann, M. R. & Walker, J. (2005). Competence and performance in causal learning. *Learning & Behavior*, 33, 211-229.
- 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.
- 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. & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. London: Basil Blackwell.
- Wasserman, E. A. (1990). Detecting response-outcome relations: Toward an understanding of the causal texture of the environment. In G. H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 26. New York: Academic Press.
- Wasserman, E. A. & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgment: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, 51B, 121-138.
- Wasserman, E. A. & Neunaber, D. J. (1986). College students' responding to and rating of contingency relations: The role of temporal contiguity. *Journal of Experimental Analysis of Behavior*, 46, 15-35.
- Wasserman, E. A., Chantosh, D. L. & Neunaber, D. J. (1983). Perception of causal relations in humans: Factors affecting judgments of response-outcome contingencies under free-operant procedures. *Learning and motivation*, 14, 406-432.
- Wasserman, E. A., Dorner, W. W. & Kao, S. F. (1990). Contributions of specific cell information to judgments of interevent contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 509-521.
- Wasserman, E. A., Eble, S. M., Chantosh, D. L. & Baker, A. G. (1993). Rating causal relations: 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 & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 34. *Causal learning* (pp. 47-88). San Diego, CA: Academic Press.
- Weiner, B. (1985). Spontaneous causal thinking. *Psychological Bulletin*, 97, 74-84.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.
- 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.
- White, P. A. (2002). Causal attribution from covariation information: The evidential evaluation model. *European Journal of Social Psychology*, 32, 667-684.
- White, P. A. (2003). Making causal judgments from the proportion of confirming instances: The pCI rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 710-727.
- Williams, D. A., Sagvoss, K. E. & McPhee, J. E. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 694-709.
- Wilson, R. A. & Keil, F. C. (1998). The shadows and shadings of explanation. *Mind and Medicine*, 8, 137-159.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- 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.
- Yates, A. S., Cheng, P. W. & Holyoak, K. J. (1995). Alternative approaches to causal induction: The probabilistic contrast versus the Rescorla-Wagner model. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 425-430). Mahwah, NJ: Erlbaum.