

Gibt es einen einheitlichen Kausalbegriff?

Andreas Bartels

Studium generale, Universität Mainz, 22.6.2005

1. Kausale Intuitionen

Fundamentale Begriffe sind fast unvermeidlich mit verschiedenen, manchmal miteinander unvereinbaren Intuitionen verbunden. Es ist dann immer die Frage, *welche* dieser Intuitionen wir in einer Begriffsdefinition auszeichnen, und welche wir zurückweisen oder einschränken. Wenn auseinander strebende Intuitionen zugrunde liegen, ist die *Einheit* eines Begriffs jedenfalls problematisch; vielleicht können wir sie garantieren, indem wir uns zwischen den verschiedenen möglichen Wegen entscheiden oder sogar eine Basis finden, auf der die *prima facie* unterschiedlichen Intuitionen sich als vereinbar oder gar sich gegenseitig ergänzend herausstellen. Vielleicht stellt es sich aber auch heraus, dass *keine* begriffliche Einheit herstellbar ist, und ein und dasselbe Wort, beispielsweise das Wort „Ursache“ der Sache nach sehr Verschiedenes bezeichnet. Für den Begriff der „Verursachung“ oder der „Ursache“ sind wenigstens die folgenden verschiedenen Intuitionen in Kraft:

REGULARITÄTEN: *Wirkungen sind regelmäßig eintretende Folgen ihrer Ursachen*

WAHRSCHEINLICHKEITEN: *Ursachen erhöhen die Wahrscheinlichkeit ihrer Wirkungen*

KONTRAFAKTISCHE KONDITIONALE: *Hätte die Ursache nicht stattgefunden, so wäre auch die Wirkung nicht eingetreten*

PROZESSE: *Ursachen übertragen physikalische Erhaltungsgrößen auf ihre Wirkungen entlang kontinuierlicher raumzeitlicher Wege*

INTERVENTIONEN: *Würde durch eine Intervention der Wert von X (Ursache) verändert, so würde sich auch der Wert von Y (Wirkung) verändern*

Evidenzen für die Intuition der Regularität sind Beispiele wie jenes des Griffs an die heiße Herdplatte, der regelmäßig ein Verbrennen der Hand zur Folge hat. Die meisten Verursachungs-Beziehungen lernen wir durch die Erfahrung regelmäßiger Abfolgen von Ereignissen kennen, so wie dies David Hume in seiner berühmten Definition der „Ursache“ ausgedrückt hat:

„We may define a cause to be an object, followed by another, and where all the objects similar to the first, are followed by objects similar to the second“ (David Hume, Enquiry concerning Human Understanding 1748, sec.VII.).

Aber auch Beispiele für die Wahrscheinlichkeits-Intuition sind nicht schwer zu finden: Rauchen ist in dem Sinne eine Ursache für Lungenkrebs, dass Rauchen die *Wahrscheinlichkeit erhöht*, an Lungenkrebs zu erkranken. Wir akzeptieren in vielen Fällen, dass Ereignisse oder Tatsachen Ursachen sind, wenn ihre Wirkungen nicht in *jedem Fall (ausnahmslos)*, aber doch immerhin häufiger auftreten als dies bei Fehlen der Ursache der Fall ist. Ein typischer Fall für die Intuition, nach der die Relation der Verursachung durch *kontrafaktische Konditionale* ausgedrückt wird, liegt in folgendem Beispiel vor: Der Schuss des Attentäters wurde vom Tod des Diktators gefolgt, eine kausale Beziehung liegt aber nur dann vor, wenn folgendes gilt: Hätte der Attentäter nicht geschossen, so wäre der Diktator nicht getötet worden. Die kausale *Prozess-Intuition* wird beispielsweise dann bestätigt, wenn eine Billardkugel beim Stoß ihre Bewegungsenergie auf die gestoßene Kugel überträgt: Um von „Verursachung“ sprechen zu können, sollte der Vorgang raumzeitlich kontinuierlich sein, und der Gesamtimpuls sollte erhalten bleiben (z.B. würden wir mit dem Attribut „kausal“ vorsichtig umgehen, wenn eine rote Kugel auf eine ruhende schwarze Kugel zurollt, nach der Kollision aber beide Kugeln in Ruhe verharren). Schließlich kann ein landwirtschaftliches Beispiel die Interventions-Intuition verdeutlichen. Dass ein Aufbringen des Düngers in der Konzentration X sich nach der Beziehung $Y = 2X$ linear in der Erhöhung des Ernteertrages niederschlägt, kann danach so verstanden werden: Würde eine Intervention erfolgen, die darin besteht, dass Dünger in der Konzentration x auf dem Feld aufgebracht wird, so würde sich dies in einer Erhöhung des Ernteertrages auf den Wert $Y = 2x$ bemerkbar machen.

Natürlich sind mit manchen der Beispiele auch mehrere kausale Intuitionen verträglich. Beispielsweise ist der Stoß von Billardkugeln auch ein Beispiel regelmäßiger Abfolge, der Tod des Diktators könnte auch probabilistisch verstanden werden: Die Abgabe des Schusses erhöht nämlich die Wahrscheinlichkeit des Todes des Diktators. Es gibt allerdings unter den genannten auch Beispiele, auf die *einige*, nicht jedoch *alle* Intuitionen zutreffen, die also in manchen Hinsichten „kausal“ sind, in anderen aber nicht. So ist etwa des Rauchen und Lungenkrebs-Beispiel, das die probabilistische Intuition veranschaulicht, kein Beispiel von Regelmäßigkeit, kein Beispiel der kontrafaktischen Intuition und jedenfalls kein klares Beispiel für die Prozess-Intuition.

Dies könnte man, sicher noch etwas voreilig, als Anzeichen dafür interpretieren, dass „Ursache“ und „Verursachung“ Begriffe im Sinne von Wittgensteins Familienähnlichkeit sind. So wie jedes Spiel in irgendeiner Hinsicht ein Spiel ist, ohne es in jeder möglichen Hinsicht zu sein, und es keine Menge von hinreichenden und notwendigen Eigenschaften gibt, die etwas zu einem „Spiel“ macht, so könnten auch Ursachen jeweils in einigen Hinsichten Ursachen sein, ohne es in allen Hinsichten sein zu müssen, und es könnte sich herausstellen, dass es kein für die Gesamtheit der Ursachen maßgebliches Kriterium gibt.

2. Grenzen der kausalen Intuitionen

Wie gesagt, dieser Schluss wäre etwas voreilig; tatsächlich können wir erst dann Bestimmteres sagen, wenn die kausalen Intuitionen zu Theorien präzisiert sind – für Zwecke unserer Betrachtung verstehen wir eine Theorie als ein Instrument, das Anwendungsfälle von „Ursache“ oder „Verursachung“ zu detektieren erlaubt. Im Einzelnen wird eine Theorie der Kausalität eine Definition von „A verursacht B“ enthalten, Anwendungsbedingungen für diese Definition festlegen, formale Eigenschaften der Verursachungsrelation ableitbar machen und Beziehungen zu verwandten Begriffen klären.

Präzisieren wir die kausalen Intuitionen nun in Form von Theorien der Kausalität, so stellt sich heraus, dass jede der einschlägigen, auf jeweils eine der genannten Intuitionen gestützten Theorien in spezifischen Situationen in ihrer Detektor-Funktion versagt, weil sie entweder bei intuitiv klaren Fällen von Verursachung ein Negativ-Resultat liefert, oder umgekehrt auf intuitiv klare Fälle von Nicht-Verursachung anspricht.

Die Anwendungsgrenze der *Regularitäts-Intuition* ist uns schon präsent: Die meisten Ursachen werden nicht ausnahmslos von ihren Wirkungen gefolgt. Deshalb hat J.L. Mackie Hume's Theorie so verfeinert, dass von Ursachen nur noch gefordert wird, nicht-hinreichende, wenn auch nicht-redundante, Bestandteile einer nicht-notwendigen aber hinreichenden Bedingung zu sein; kurz: Ursachen sind *INUS-Bedingungen*, unverzichtbare Bestandteile eines zusammen für die Wirkung hinreichenden Bedingungskomplexes – wobei in der Regel mehr als ein solcher Bedingungskomplex, zusammengesetzt aus jeweils verschiedenen Bestandteilen, existiert. Wenn im Sinne dieser Theorie festgestellt wird, dass das Entzünden eines Streichholzes den Waldbrand verursacht hat, so wird nicht mehr wie in der ursprünglichen Regularitätstheorie gefordert, dass das Anzünden von Streichhölzern automatisch Waldbrände zur Folge hat; es ist eben nicht *eine* Ursache, sondern ein ganzer *Bedingungskomplex*, der hinreichend für die Wirkung sein soll. Das Problem, das solche

Gegenevidenzen wie der Fall „Rauchen verursacht Lungenkrebs“ der Regularitätstheorie immer bereitet haben, bleibt allerdings erhalten: Denn wir beobachten auch nicht irgendwelche Bedingungskomplexe, bei dessen Auftreten unweigerlich Lungenkrebs die Folge ist (was andererseits gut „passt“ ist, dass Lungenkrebs auch ganz ohne Rauchen auftreten kann, dann also offenbar durch einen anderen Bedingungskomplex herbeigeführt, der das Rauchen nicht als Bestandteil enthält). Ein weiteres notorisches Problem jeder Regularitätstheorie ist, dass sie „unechte“ Regularitäten nicht erkennt. Das Absinken des atmosphärischen Drucks verursacht einerseits, dass „das Barometer fällt“, andererseits, dass sich etwas später ein Sturm einstellt. Die gemeinsame Ursache „Absinken des atmosphärischen Drucks“ führt daher zu einer regelmäßigen Korrelation der von ihr gemeinsam verursachten, kausal voneinander völlig unabhängigen Wirkungen. Aber aufgrund der Regularitätstheorie können wir dies nicht erkennen: sie wird diese Korrelation als kausal auszeichnen. Wenn wir „mit bloßem Auge“ erkennen, dass es hier nicht um Verursachung geht, so offenbar indem wir die „Prozess-Intuition“ hinzunehmen: Wir glauben, dass es keinen kausalen Prozess gibt, der den Barometerstand (direkt) mit dem Sturm verbindet.

Diese beiden Sorten von Gegenevidenzen der Regularitätstheorie lassen folgenden Schluss zu: Die Regularitätstheorie scheitert daran, dass sie zwei Tatsachen unberücksichtigt lässt:

- (1) Viele Ursachen erhöhen nur die Wahrscheinlichkeit ihrer Wirkungen, ohne *hinreichend* für sie zu sein, und
- (2) Es gibt Regularitäten (Korrelationen), die *nicht kausal* sind, sondern durch gemeinsame kausale Hintergrund-Faktoren (common causes) erzeugt werden.

Diesen beiden Tatsachen wird die *Probabilistische Theorie der Kausalität* gerecht:

A verursacht B dann und nur dann, wenn für die bedingte Wahrscheinlichkeit $P(B/A)$ gilt:
 $P(B/A) > P(B/\text{nicht } A)$ und es existiert kein Abschirmfaktor C mit $P(B/A \text{ und } C) = P(B/C)$.

Die Existenz eines Abschirmfaktors bedeutet, dass die Korrelation zwischen A und B auf einen gemeinsamen kausalen Faktor C zurückgeführt werden kann, so wie dies im Barometer-Beispiel der Fall ist: Nimmt der atmosphärische Druck C einen bestimmten Wert an, so wird die Wahrscheinlichkeit des Auftretens eines Sturmes B jedenfalls nicht dadurch verändert, dass das Barometers einen Wert A anzeigt. Wenn ein Abschirmfaktor C existiert, auf den die Korrelation zwischen A und B zurückgeführt werden kann, so spricht diese Korrelation nicht für eine Kausalbeziehung zwischen A und B. Dass Korrelationen aufgrund der Existenz von

common causes von „echten“ Kausalbeziehungen unterschieden werden, zeigt, dass neben der probabilistischen Intuition eine weitere, die kausale Prozessintuition im Spiel ist; die Motivation der Abschirmbedingung ist es, solche probabilistischen Beziehungen, die auf echten kausalen Pfaden zustande kommen, von jenen zu unterscheiden, die dazu parasitär sind.

Leider gibt es unechte Korrelationen, die nicht durch die Abschirmbedingung erfasst werden. Es gibt Ursachen, die die Wahrscheinlichkeit des Eintretens ihrer Wirkung *verringern*. So haben Nichtraucher eine höhere Wahrscheinlichkeit, einen Herzinfarkt zu erleiden (B), als Raucher (A), weil unter Rauchern die Anzahl der Sport treibenden Menschen (C) signifikant höher ist als unter Nichtrauchern, obwohl in den Bezugsgruppen der Sport treibenden wie in jener der nicht Sport treibenden das Rauchen jeweils die Wahrscheinlichkeit des Herzinfarktes erhöht:

$P(B/A) < P(B/\text{nicht } A)$, obwohl
 $P(B/A \text{ und } C) > P(B/\text{nicht } A \text{ und } C)$ sowie
 $P(B/A \text{ und nicht } C) > P(B/\text{nicht } A \text{ und nicht } C)$

Diese paradox erscheinenden Korrelationen lassen sich damit erklären, dass C ein hoch signifikanter Faktor der Verhinderung von B ist, so dass die größere Häufigkeit des Auftretens von C in der Gruppe der Raucher gegenüber der Gruppe der Nicht-Raucher die Wirksamkeit des Risikofaktors „Rauchen“ kompensiert.

Erhöhung oder Verringerung der Wahrscheinlichkeit kann bestenfalls *ein* Indikator für das Vorliegen von Kausalbeziehungen sein. Wenn ich Pech habe, verursache ich gerade dadurch den Tod der alten Dame, dass ich sie vor dem herannahenden Bus von der Straße auf den Bürgersteig ziehe, weil ihr dann unglücklicherweise auf dem Bürgersteig ein Dachziegel (ein sehr unwahrscheinliches Ereignis!) auf den Kopf fällt. Ich verursache ihren Tod, obgleich ich ihre Wahrscheinlichkeit zu sterben stark *verringert* habe. Weder erscheint es plausibel, in einem solchen Fall *nicht* von Verursachung zu sprechen, noch kann die Berechtigung von Verursachung zu sprechen, vom Verhältnis der Wahrscheinlichkeiten, durch einen Bus bzw. durch einen Dachziegel zu sterben, abhängen. Der Wahrscheinlichkeits-Vergleich darf nicht zwischen den beiden miteinander verschränkten Prozessen vorgenommen werden, sondern zwischen jeweils *einem* der Prozesse und einem alternativen, *kontrafaktischen* Prozess: Durch

meine Aktion habe ich den Tod der alten Dame *verhindert* (bzw. im Fall des Scheiterns immerhin *gehindert*), weil durch sie ihre Wahrscheinlichkeit durch den Bus zu Tode zu kommen, verringert worden ist, *verglichen mit der Situation, in der meine Aktion ausgeblieben wäre*. Kausale Aussagen sind eben auf bestimmte kausale Pfade zu spezifizieren und hinsichtlich dieser spezifischen Pfade durch ein kontrafaktisches Argument zu erläutern. Daraus folgt dann allerdings auch, dass die Bedeutung kausaler Aussagen auf den Kontext eines kausalen Prozesses eingeschränkt wird: Ich verhinderte den Tod der alten Dame – aber nur in Hinsicht auf die Verminderung ihrer Chance durch den nahenden Bus zu sterben – gleichwohl verursachte ich mit derselben Aktion ihren Tod durch den herabfallenden Dachziegel, weil diese Aktion ihre Chance der Kollision mit dem Dachziegel erhöhte. Verursachung ist, wie Phil Dowe es formuliert hat, „where an event raises the chance of another relative to genuine causal processes linking the two, and the effect occurs” (Phil Dowe: “Promoting, Causing, Hindering, Preventing”. In: Spohn/Ledwig/Esfeld 2001, S. 81).

Beispiele wie diese zwingen uns dazu, die ursprüngliche naive Variante der probabilistischen Theorie zu modifizieren: Ursachen müssen die Wahrscheinlichkeit ihrer Wirkungen *in Testsituationen* bzw. hinsichtlich raumzeitlich spezifizierter *kausaler Prozesse* erhöhen. In Testsituationen werden alle von der betrachteten potentiellen Ursache *unabhängigen Ursachen festgehalten*. Es gibt eben in vielen Situationen mehrere kausale Einflussfaktoren, die möglicherweise gegenläufige Einflüsse ausüben. Um die Wirkung eines Faktors zu isolieren, müssen daher die anderen Faktoren in ihrer Wirkung eliminiert werden. Wieder scheint die kausale Prozessintuition eine Rolle zu spielen: Um die statistischen Daten „zum Sprechen zu bringen“, müssen aus einem Netz von Kausalbeziehungen *einzelne kausale Pfade* isoliert werden. Und die Methode, um einzelne kausale Pfade zu isolieren, ist eine *kontrafaktische Analyse*: „A verursacht B“ bedeutet: A würde die Wahrscheinlichkeit von B erhöhen, wenn die unabhängigen kausalen Faktoren C festgehalten würden... Wir sehen im Rahmen des probabilistischen Konzepts, dass die Bedeutung kausaler Aussagen ohne einen Bezug auf kontrafaktische Situationen nicht explizierbar ist. Dies bringt mit sich, dass wir, um über Verursachung zwischen A und B sprechen zu können, schon über *andere kausale Beziehungen*, nämlich jene zwischen C und B... sprechen müssen; kausales Vokabular kann offenbar nicht auf nicht-kausales Vokabular reduziert werden: *no causes in, no causes out*.

Die Probleme einer reinen probabilistischen Auffassung führen uns also die Berechtigung kontrafaktischer Intuitionen der Verursachung vor Augen, die den Hintergrund für David

Lewis' *kontrafaktische Theorie* bilden. Lewis hat darauf aufmerksam gemacht, dass der zweite Teil der vorhin im Zusammenhang mit der Regularitätstheorie zitierten Äußerung von David Hume keineswegs eine Paraphrasierung des ersten Teils, sondern die erste Formulierung einer kontrafaktischen Kausalitätskonzeption darstellt:

„We may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed” (Hume 1748, section VII).

Dass dieser Gedanke so lange auf seine Ausführung warten musste, hat vor allem damit zu tun, dass Wahrheitsbedingungen für kontrafaktische Sätze nicht leicht anzugeben sind: Unter welchen Umständen ist ein kontrafaktischer Satz der Art *„Wäre auf Höhe der Torlinie eine Kamera installiert gewesen, so wäre das „Wembley-Tor“ nicht anerkannt worden“* wahr? Klar ist zunächst nur, dass ein kontrafaktisches Konditional nicht wahrheitsfunktional ist, also seine Wahrheit oder Falschheit nicht allein von der Wahrheit oder Falschheit der Teilsätze (*Auf Höhe der Torlinie war eine Kamera installiert* bzw. *Das Wembley-Tor wurde nicht anerkannt*) abhängt. In unserem Beispiel sind beide falsch, aber ob der kontrafaktische Konditionalsatz wahr oder falsch ist, lässt sich daraus nicht erschließen. David Lewis hat dieses Problem gelöst, indem er auch Tatsachen in anderen möglichen Welten mit ins Kalkül gezogen hat. In der Tat meinen wir mit dem kontrafaktischen Konditionalsatz so etwas wie: *Wenn die Welt nur um einen Hauch anders gewesen wäre, nämlich so, dass die Kamera installiert gewesen wäre, während alles Übrige gleich gewesen wäre, dann wäre das Wembley-Tor nicht anerkannt worden.* Anerkennen oder Nicht-Anerkennen des Tores hing gewissermaßen nur an diesem einen dünnen Faden, der Tatsache, dass eine Kamera installiert oder nicht installiert ist. In der Terminologie der möglichen Welten: Es gibt eine mögliche Welt, in der – abweichend von der wirklichen Welt – die Kamera installiert ist, und das Wembley-Tor nicht anerkannt wird, und diese Welt ist unserer Welt so ähnlich wie nur irgend möglich (sie unterscheidet sich nur in Hinsicht auf diese Kamera). Jedenfalls ist jede andere mögliche Welt, in der die Kamera installiert ist und das Tor anerkannt wird, der wirklichen Welt *weniger ähnlich*, d.h. sie unterscheidet sich von der wirklichen Welt in mehr Tatsachen als nur jener in Hinsicht auf die Kamera.

Lewis hat diese Überlegung in folgende Formel gegossen:

Lewis (1973): $A \square \rightarrow C$ ist wahr in der aktualen Welt genau dann wenn (i) irgendeine A-Welt, in der C gilt, näher an der aktualen Welt ist als alle A-Welten, in denen C nicht gilt, oder wenn (ii) es keine möglichen A-Welten gibt.

Dass E von C *kausal abhängt*, bedeutet nun nach Lewis, dass das folgende kontrafaktische Konditional wahr ist: *Hätte C nicht stattgefunden, so hätte auch E nicht stattgefunden*. C ist dann eine *Ursache* von E (nicht notwendig *die* Ursache), wenn es eine Kette kausaler Abhängigkeit gibt, die C mit E verbindet (diese Festlegung macht Verursachung transitiv). Aber auch Lewis' Theorie hat mit widerspenstigen Beispielen zu kämpfen: Nehmen wir an, in unserem Diktator-Beispiel gebe es zwei Attentäter A und B; es ist A, der zuerst abdrückt, worauf B den Schuss unterlässt. Nun gilt nicht mehr: „Hätte Attentäter A nicht geschossen, wäre der Diktator noch am Leben“. Hätte nämlich A nicht geschossen, so hätte B geschossen. Hier kann man der Verursachungs-Relation zwischen dem Schuss und dem Tod des Diktators durch Einschleiben eines intermediären Ereignisses gerecht werden: Wir betrachten ein Ereignis F, das darin besteht, dass die Kugel irgendwo zwischen A und dem Diktator auf dem Weg ist. Dann hängt der Tod des Diktators kausal von F ab (...hätte F nicht stattgefunden, wäre der Diktator tatsächlich noch am Leben...), und F hängt kausal vom Abdrücken durch A ab, also ist das Abdrücken die Ursache des Todes.

Schwieriger wird es in Fällen des *kausalen Zuvorkommens* (*causal preemption*). Jetzt schießen A und B, aber es ist A's Kugel, die den Diktator tötet. Lewis' Theorie erklärt nicht, dass A's Schuss die Ursache des Todes war, denn hätte A nicht geschossen, hätte B's Kugel den Diktator getötet (und für intermediäre Ereignisse gilt analoges). „Intuitiv“ ist A's Schuss die Ursache des Todes, aber nach Lewis' Theorie ist der Schuss keine Ursache des Todes. Ein möglicher Ausweg ist, den Fall einmal so zu betrachten, als wäre B gar nicht anwesend gewesen. Wieder „intuitiv“ gesprochen, wäre dann alles kausal Relevante genauso gewesen wie im ursprünglichen Beispiel – und es wäre ganz unzweifelhaft nach Lewis' Theorie so, dass A's Schuss den Tod des Diktators verursacht. Da die beiden Fälle durch nichts kausal Relevantes unterschieden sind, muss diese Aussage also auch für das ursprüngliche Beispiel gelten. Ich habe nichts gegen diese Lösung einzuwenden, aber vielleicht haben auch Sie den Eindruck, dass die Lösung doch nur deswegen plausibel erscheint, weil die Intuition, Verursachung habe etwas mit raumzeitlich kontinuierlichen Prozessen zu tun, zu Hilfe genommen wird. Dass der alternative Fall (ohne B) dem ursprünglichen in allem kausal Relevanten gleicht, gilt ja nur dann, wenn für die Verursachung des Todes durch den Schuss letztlich nur der von A ausgelöste physikalische Prozess entscheidend ist.

Diese Intuition taucht also im Rahmen eines kontrafaktischen Konzeptes geradezu zwangsläufig auf. Über Bertrand Russell, Hans Reichenbach, Wesley Salmon, David Fair bis zu Phil Dowe reicht die Tradition der *kausalen Prozess-Theorie*. Die entscheidende Aufgabe einer solchen Theorie ist es, ein Unterscheidungsmerkmal anzugeben, das kausale Prozesse (z.B. die Ausbreitung elektromagnetischer Signale) von nicht-kausalen Abfolgen von Zuständen (z.B. der Persistenz eines Schattens) unterscheidet; dafür wurden eine Reihe von Lösungen vorgeschlagen; in der gegenwärtig diskutierten Form der Theorie sind es physikalische Erhaltungsgrößen (Energie-Impuls), die ein kausaler Prozess trägt (oder der in einem Prozess übertragen wird).

Was immer aus theoretischer Sicht für das Konzept spricht, es stellt sich ein grundlegendes Anwendungsproblem. Selbst wenn alle Verursachungen in unserer Welt, mikroskopisch betrachtet, physikalische Prozesse sind, die Energie-Impuls übertragen, hilft uns dies nicht bei der Beantwortung der Frage, welche der zahlreichen statistischen Abhängigkeiten in einem gegebenen Feld von Variablen als *kausale Abhängigkeiten* zu interpretieren sind. Um herauszufinden, ob tatsächlich Rauchen Lungenkrebs verursacht, ist es irrelevant, dass kausale Prozesse stattgefunden haben, die Ereignisse des Rauchens einer Person mit Ereignissen des Auftretens von Krebszellen verbinden. Das kausale Prozess-Konzept enthält keine Information darüber, *welche* Eigenschaften in *kausalen* Beziehungen zueinander stehen. Wir erfahren etwas über das *ontische Substrat von Verursachung*, aber nichts über die *kausale Struktur der Welt*, soweit es sich um Beziehungen zwischen Eigenschaften handelt.

Das Zwischenresultat unserer Überlegungen fällt ernüchternd aus: Auf der einen Seite ist keine der etablierten Theorien der Kausalität, Regularitätstheorie, probabilistische Theorie, kontrafaktische Theorie für sich genommen in der Lage, die Gesamtheit unserer kausalen Urteile verständlich zu machen, auf der anderen Seite ist die kausale Prozesstheorie, die deren Defizite zu beheben verspricht, wenig hilfreich, wenn es darum geht zu entscheiden, ob Korrelationen zwischen makroskopischen, ausgedehnten Eigenschaften (wie „Rauchen“ und „Lungenkrebs“) kausal oder nicht kausal zu interpretieren sind. Im zweiten Teil des Vortrags werde ich daher auf das *Interventions-Konzept* eingehen, das durch die Praxis kausalen Modellierens motiviert ist. Dieses Konzept greift die grundlegende Intuition der kontrafaktischen Kausalitätsauffassung auf: Kausale Aussagen sind kontrafaktische Aussagen darüber, was unter hypothetischen, idealisierten experimentellen Manipulationen (an einer unabhängigen Variablen) geschehen würde (d.h. wie eine abhängige Variable durch diese

Manipulation beeinflusst würde). Zugleich wird die Intuition kausaler Prozesse nutzbar gemacht: Originäre (direkte) kausale Beziehungen (symbolisiert durch *kausale Pfade*) hängen nur von den intrinsischen Eigenschaften ihrer Relata ab und nicht von der kausalen Umgebung. Kausale Beziehungen können stets auf *direkte* kausale Beziehungen zurückgeführt werden und es ist das Ziel der kausalen Modellierung, die Struktur der direkten kausalen Beziehungen zu finden, die den Beobachtungsdaten (statistische Daten) zugrunde liegen.

3. Kausale Interventionstheorie

Das Basismaterial jeder kausalen Modellierung sind statistische Daten, die Abhängigkeiten zwischen einer Reihe von Variablen ausdrücken, wie in folgendem Beispiel:

$$\begin{aligned} X_2 &= a X_1 \\ X_3 &= b X_1 \\ X_4 &= c X_2 + d X_3 \end{aligned}$$

Diese Daten werden als *Hinweise* auf die gesuchten kausalen Pfade gelesen. Kausales Modellieren besteht darin, die echten kausalen Pfade aus den Daten zu extrahieren. Idealerweise stehen im obigen Gleichungssystem jeweils auf der rechten Seite die direkten kausalen Ursachen der auf der linken Seite stehenden Variablen. Das Interventions-Konzept stellt nun ein Instrument zur Verfügung, um zu überprüfen, ob diese ideale Situation tatsächlich gegeben ist. Wäre tatsächlich X_1 die direkte kausale Ursache von X_2 mit dem durch den Koeffizienten a angegebenen Einfluss, so müsste sich nach diesem Konzept eine Intervention hinsichtlich von X_1 genau in der angegebenen Weise in einer Veränderung von X_2 niederschlagen. Diese Abhängigkeit kann im Prinzip experimentell getestet werden.

Wenn wir nun dasselbe für die Beziehung zwischen X_2 und X_4 durchführen wollen, stoßen wir auf das Problem der Überlagerung von Wirkungen: Änderungen von X_2 , die durch Interventionen hervorgerufen werden, die sich auch in X_1 niederschlagen, würden sich (auf dem Weg $X_1 \rightarrow X_3 \rightarrow X_4$) ebenso in X_4 bemerkbar machen wie Wirkungen, die auf direktem Wege von X_2 nach X_4 verlaufen. Die im Test erhaltenen Daten würden dann eben keinen verlässlichen Rückschluss auf die direkte kausale Abhängigkeit zwischen X_2 und X_4 erlauben. Um festzustellen, ob X_2 einen direkten kausalen Einfluss auf X_4 ausübt, müssen wir die *indirekten* Wirkungen auf X_4 eliminieren, d.h. Interventionen hinsichtlich X_2 , die den Einfluss

auf X_4 testen, sollten nicht mit anderen kausalen Faktoren in der Umgebung interferieren. Dies kann durch folgende Bedingungen an die Intervention I erreicht werden:

- (1) Die Intervention I verursacht X_2 .
- (2) I hat keinen Einfluss auf Variable außer jenen, die auf dem direkten kausalen Pfad von X_2 nach X_4 liegen.
- (3) I ist probabilistisch unabhängig von jeder anderen Variablen, die kausal relevant für X_4 ist (z.B. X_1).
- (4) I verändert nicht die Beziehungen zwischen X_4 und seinen Ursachen außerhalb des direkten kausalen Pfades.
- (5) I setzt jeden Einfluss der Eltern von X_2 außer Kraft (I annulliert die erste Gleichung des Gleichungssystems).
- (6) I ändert nicht die kausalen Beziehungen zwischen X_2 und X_4 (*Invarianz-Forderung*).

Diese Bedingungen legen fest, was unter einer *Intervention* im Sinne der Theorie zu verstehen ist, und damit, was es heißt, statistische Abhängigkeiten zwischen Variablen als Repräsentation direkter kausaler Beziehungen zu interpretieren. Statistische Daten im Sinne des Gleichungssystems oben sind eben nur dann im Sinne eines kausalen Pfadmodells zu interpretieren, wenn Interventionen hinsichtlich der unabhängigen Variablen (linke Seite des Gleichungssystems) diese Daten bestätigen. Daraus folgt auch, dass nicht jede Darstellung der statistischen Daten als Repräsentation der *kausalen* Beziehungen interpretiert werden kann. In unserem Beispiel könnten wir etwa in der letzten Gleichung X_2 und X_3 durch X_1 ausdrücken: $X_4 = (ca+db) X_1$. Die Gleichungssysteme sind hinsichtlich der Beobachtung ununterscheidbar, sie repräsentieren genau dieselbe Menge statistischer Daten. Die neue Gleichung verrät aber nichts mehr über den direkten kausalen Einfluss der Variablen X_2 und X_3 auf X_4 , d.h. sie ist keine korrekte Repräsentation der kausalen Beziehungen des Systems.

Der Forderungskatalog impliziert, dass „Intervention“ selbst schon ein kausaler Begriff ist; so wird ja z.B. gefordert, dass I X_2 *verursacht*, dass I nicht die Beziehungen zwischen X_4 und seinen *Ursachen* außerhalb des direkten kausalen Pfades verändert u.s.w. Die Interventionstheorie verheißt also keine Reduktion kausaler auf nicht-kausale Begriffe.

Ein naheliegender Einwand gegen das Interventionskonzept ist, dass in vielen Fällen ideale Interventionen nicht ausführbar sind. Um Placebo-Wirkungen von Medikamenten zu eliminieren, kann man aus ethischen Gründen das Medikament nicht einfach ohne Wissen des

Patienten verabreichen, um so den über den Placebo-Faktor laufenden kausalen Pfad abzuschneiden. Tatsächlich wird in einem solchen Fall so vorgegangen, dass sowohl die Behandlungs- wie die Kontrollgruppe einem Placebo-Effekt ausgesetzt wird (die Kontrollgruppe enthält ein unwirksames Medikament). Auf diese Weise kann der Placebo-Effekt, der die Behandlungsgruppe betrifft, „herausgerechnet“ werden; wir können das Verfahren also so verstehen, dass es darauf abzielt zu ermitteln, was geschehen *würde*, wenn die Intervention (Verabreichung des Medikamentes) ohne Wissen der Patienten vorgenommen würde. Bei näherem Hinsehen stellt sich also heraus, dass das praktische Vorgehen ganz im Sinne der Interventionstheorie verstanden werden kann.

Eine wichtige, aber umstrittene Forderung an Interventionen ist die *Modularitäts-Bedingung*. Sie besagt, dass wenigstens für einen bestimmten Wertebereich die Gleichungen eines kausalen Systems invariant gegenüber Interventionen hinsichtlich der unabhängigen Variablen des Systems sind. Man kann dies auch so ausdrücken, dass es möglich ist eine einzelne Gleichung des Systems zu „annulieren“ (indem die abhängige Variable durch eine Intervention mit einem bestimmten Wert ersetzt wird), ohne damit die anderen Gleichungen des Systems anzutasten. Intuitiv bedeutet dies, dass die einzelnen kausalen Mechanismen des Systems unabhängig voneinander arbeiten. Wir können in einen der Mechanismen eingreifen, ohne damit notwendigerweise die anderen Mechanismen zu stören.

Am Barometer-Beispiel kann man einsehen, wie wichtig diese Forderung ist. Die maßgeblichen Gleichungen seien hier (S = Sturm, A = Atmosphärischer Druck, B = Barometer-Stand): $S = a A$ und $B = b A$. Die Modularitätsbedingung besagt nun, dass wir bei Unterstellung einer common-cause-Struktur durch eine Intervention die Beziehung $A \rightarrow B$ zwischen dem atmosphärischen Druck und dem Barometerstand „annulieren“ können, indem wir den Zeiger des Barometers gewaltsam auf einen bestimmten Wert einstellen, ohne damit die Beziehung $A \rightarrow S$ zwischen dem atmosphärischen Druck und dem Auftreten eines Sturmes zu stören. Würde sich dies nicht bestätigen (würde das Einstellen des Barometers mit der Neigung zum Auftreten eines Sturms korrelieren), so müssten wir die Annahme der common-cause-Struktur aufgeben. Stattdessen würden wir in einem solchen Fall vermuten, dass doch ein direkter kausaler Einfluss zwischen dem Barometerstand und dem Eintreten eines Sturmes besteht. Fälle, in denen die Modularitäts-Bedingung nicht erfüllt ist, wären also zugleich Falsifikatoren unseres kausalen Modells.

Die Interventionstheorie stellt eine Synthese der bisher dargestellten Konzepte der Kausalität dar. Als das hartnäckigste Problem des rein probabilistischen Konzeptes hatten sich wahrscheinlichkeitsvermindernde Verursachungen herausgestellt. Um es zu beheben, müssen wir die kausale Umgebung eines Prozesses „ausblenden“ – und den probabilistischen Effekt, den der untersuchte Faktor auf seinem kausalen Weg, auf den wir den Focus richten, mit jenem vergleichen, der auf dem gleichen Weg bei Fehlen des Faktors auftreten würde. Wir benötigen also ein *kontrafaktisches Konditional* und die Annahme der *relativen Unabhängigkeit der kausalen Mechanismen*. Das hartnäckigste Problem des kontrafaktischen Konzepts selbst stellen, wie gesehen, Fälle zuvorkommender Verursachung dar. Auch hier benötigen wir zur Lösung des Problems eine Entwirrung der kausalen Pfade, d.h. das Ausblenden der kausalen Umgebung eines Prozesses, um zu garantieren, dass das kontrafaktische Konditional, das sich auf den betrachteten, aus seiner kausalen Umgebung herausgehobenen Prozess bezieht, dessen kausale Natur ausdrückt.

Die Mittel zur Lösung der in den „reinen“ Kausalitätstheorien ungelösten Probleme lassen sich daher allesamt auch in Begriffen des interventionistischen Konzepts ausdrücken: Die kontrafaktischen Konditionale „Hätte C nicht stattgefunden, dann wäre E nicht aufgetreten“ bzw. „Hätte C nicht stattgefunden, wäre die Wahrscheinlichkeit von E höher (niedriger) gewesen“ sind im Interventionskonzept repräsentierbar als: „Wäre durch eine Intervention im System C verhindert worden (z.B. der Schuss auf den Diktator), so hätte sich E (im Beispiel: der Tod des Diktators) nicht ereignet“ bzw. „Wäre durch eine Intervention im System C verhindert worden (z.B. wäre mein Eingreifen hinsichtlich der alten Dame, beispielsweise durch mein eigenes Stolpern, verhindert worden), dann wäre die Wahrscheinlichkeit des Todes der alten Dame höher gewesen“. Das andere Desiderat, die Annahme der relativen Unabhängigkeit der kausalen Mechanismen, also die Möglichkeit die kausale Umgebung „auszublenen“, gehört sogar zu den expliziten Bedingungen, die an den Begriff der Intervention gestellt werden. Daraus folgt, dass das Interventionskonzept allgemein genug ist, um notwendige Modifikationen zur Rettung der herkömmlichen Kausalkonzepte zu integrieren. In diesem Sinne erweist es sich als das umfassendste der gegenwärtig diskutierten Kausalkonzepte.

4. Einheit des Kausalbegriffs ?

Nach Nancy Cartwright ist die Einheit des Kausalbegriffs eine Chimäre. Der Begriff „Ursache“ ist, so Cartwright, eine Art Sammelbezeichnung, unter die eine Vielzahl

verschiedener Phänomene subsumiert werden, die sprachlich durch verschiedene Verben repräsentiert werden, beispielsweise „Anziehung“, „Abstoßung“, „Erhöhung“, „Erniedrigung“, „Unterbrechung“, „Hineinpumpen“, „Umstoßen“ u.s.w.:

„Causes make their effects happen. That is more than, and different from, mere association. But it need not be one single different thing. One factor can contribute to the production of another in a great variety of ways. There are standing conditions, auxiliary conditions, precipitating conditions, agents, interventions” (Cartwright 1999: Causal Diversity and the Markov Condition, *Synthese* 121: 3-27, p.18).

Hinsichtlich der Bedingung der Modularität, bemerkt Cartwright, sie charakterisiere einen bestimmten Typ kausaler Systeme, aber keineswegs *alle* kausalen Systeme. Ihr bevorzugtes Beispiel ist das des Toasters, bei dem eine kausale Beziehung zwischen dem Hebel, der bei Erreichen einer bestimmten Temperatur ausgelöst wird, und dem durch ihn hochgeschobenen Toasthalter besteht. Nun ist diese kausale Beziehung vollkommen determiniert, die beiden Teile sind fest miteinander verbunden. Also können wir diese Beziehung nicht unterbrechen, ohne die Integrität des Toasters selbst zu stören. Man könnte sagen, dass die Trennung dieser Verbindung, die es ermöglicht, den Halter frei zu verschieben, dazu führt, dass wir keinen Toaster mehr vor uns haben. Solange der Toaster intakt bleibt, müssen alle kausalen Einflüsse auf den Halter durch den Hebel ausgeübt werden; da die Auslösung des Hebels aber zugleich den Stromkreis unterbricht, gibt es keine Intervention hinsichtlich des Halters, die andere kausale Beziehungen unbeeinflusst lässt. Dem kann man entgegen halten, dass einfache Mechanismen wie die zwischen Hebel und Halter in verschiedenen technischen Geräten vorkommen – also mit ganz anderen kausalen Umgebungen als sie für den Toaster charakteristisch sind. Wir können die kausale Verbindung zwischen Hebel und Halter trennen, um zu analysieren, welche Änderungen im System durch die Verschiebung des Halters *direkt* ausgelöst werden. Es ist ja gerade eine wesentliche Voraussetzung der Ingenieurskunst, dass kausale Mechanismen unabhängig voneinander vielfach kombiniert werden können:

„One of the central ideas of the interventionist account is a „Galilean“ idea about the function of experiments: one can learn about the causal structure of a complex system by disrupting some parts of it while leaving other parts intact; by taking the system apart, trying to understand whatever principles govern its components, taken in isolation, and then understanding the overall behavior of the system as the result of the principles governing these individual components“ (Woodward: Invariance, Modularity, and All That: Cartwright on Causation).

Nun sind noch weitere Beispiele denkbar, in denen die Modularitäts-Bedingung *prima facie* nicht erfüllt ist. Die sogenannten „Spiegelneuronen“ verursachen offenbar zugleich die Wahrnehmung von bestimmten Bewegungen anderer Lebewesen wie auch die eigene Ausführung dieser Bewegungen. Ein Durchtrennen der einen kausalen Verbindung würde daher offenbar auch den anderen kausalen Mechanismus unterbrechen. In diesem Fall kann man aber sicher davon ausgehen, dass die Ausführung der einzelnen Wirkungen der Spiegelneuronen noch durch weitere neuronale Einheiten mitgesteuert wird (die Tatsache, dass die Wirkungen unabhängig voneinander beobachtet werden können, deutet darauf hin), so dass diese weiteren kausalen Bedingungen für einen Mechanismus ausgeschaltet werden können, ohne den anderen Mechanismus damit zu unterbinden.

Um es hier noch einmal zu wiederholen: Die Strategie der Modularitäts-Bedingung unterstellt, dass kausale Verbindungen zwischen den restlichen Komponenten der ursprünglichen, zu analysierenden Systems in derselben Weise weiter bestehen, wenn an einer von ihnen eingegriffen wird. Angenommen, wir besitzen eine Repräsentation der beobachtbaren Daten, durch die die *direkten Kausalbeziehungen* modelliert werden. Dann gibt es stets empirisch äquivalente Darstellungen, die dadurch zustande kommen, dass die Beiträge verschiedener direkter kausaler Pfade miteinander vermischt werden; sehen können wir dies daran, dass Unterbrechungen eines einzelnen kausalen Pfades Rückwirkungen auch für die anderen kausalen Pfade erzeugen – mithin an einer Verletzung der Modularität. Eine Verletzung der Modularität, so können wir annehmen, wird eben typischerweise durch eine Vermischung der kausalen Pfade hervorgerufen; die Modularitätsverletzung zeigt uns, dass unsere Repräsentation der kausalen Beziehungen inkorrekt ist. Positiv ausgedrückt: Ist ein kausales Modell der Daten korrekt, so erfüllt es die Modularitätsbedingung.

Trotzdem bleibt die Erfüllung der Modularität eine kontingente Eigenschaft von Systemen; sie besitzt keinesfalls apriorischen Status. Ziel einer Theorie der Kausalität kann nur die Angabe möglichst *allgemeiner*, d.h. allen einschlägigen kausalen Intuitionen

möglichst umfassend gerecht werdender Bedingungen sein, die kausale von nicht-kausalen Beziehungen unterscheiden. Ein apriorischer Kausalbegriff, der unabhängig von allen kausalen Erfahrungen universell gültige Bedingungen für die Verursachungsrelation angibt (die in den Begriff der Verursachung gleichsam „eingebaut“ wären), scheint – v.a. aufgrund der eingangs geschilderten Vielfalt kausaler Intuitionen – ausgeschlossen. Vielleicht, oder sogar sehr wahrscheinlich, ist mit der Modularitätsbedingung noch nicht die letzte Stufe der Suche nach Bedingungen erreicht, die genügend allgemein (alle kausalen Intuitionen umfassend) und doch inhaltsreich genug sind, um als Kriterium anwendbar zu sein. Aber die Einwände von Cartwright gegen eine Einheit des Kausalbegriffs lassen sich ungeachtet dessen zurückweisen, dass ein solcher perfekter Kausalbegriff noch nicht in Reichweite ist. Alle von Cartwright erwähnten kausalen Phänomene lassen sich alle unter den interventionistischen Begriff subsumieren: sie entsprechen einfach verschiedenen Formen der Abhängigkeit innerhalb eines kausalen Gleichungssystems (z.B. Anziehung und Abstoßung ebenso wie verschiedene funktionale Formen der kausalen Abhängigkeit).

Warum aber überhaupt nach einer Einheit des Kausalbegriffs streben – wenn es doch keine metaphysische Gewissheit seiner Existenz gibt? Der Versuch, Verursachung in einer allgemeinen und einheitlichen Weise zu interpretieren, bietet den Vorteil, unserer Frage danach, ob eine bestimmte Beziehung kausal ist, eine *spezifische* Bedeutung zu geben; wenn wir einen einheitlichen Begriff unterstellen, so wissen wir, wonach wir suchen, und können daher mögliche Heuristiken besser begrenzen. Dies ist gerade dort von Belang, wo erhebliche Uneinigkeit darüber besteht, welche Aussagen als kausal zu verstehen sind und welche formalen Eigenschaften kausale Beziehungen teilen (z.B. die Eigenschaft der Transitivität). Würden wir Kausalität in dieser ebenso wie in jener Bedeutung zulassen, würde die Suche nach Antworten ziellos. Und schließlich sind nur hinsichtlich eines einheitlichen Kausalbegriffs definite Antworten auf die Frage nach den Beziehungen des Kausalbegriffs zu den Begriffen seines Begriffsumfeldes zu erwarten: Erklärung, Bestätigung, Test, Experiment etc. Dies sind einige, aber vielleicht nicht alle Gründe, die für die Suche nach einem einheitlichen Kausalbegriff sprechen.

Literatur

Ernest Sosa/Michael Tooley (eds.): Causation, Oxford University Press: Oxford 1993

Wolfgang Spohn/Marion Ledwig/Michael Esfeld (eds.): Current Issues in Causation, Mentis: Paderborn 2001

Judea Pearl: Causation. Models, Reasoning, and Inference, Cambridge University Press: Cambridge 2000

Nancy Cartwright: Nature's Capacities and their Measurement, Oxford University Press: Oxford 1994

Nancy Cartwright: Causation: One World; Many Things. *Philosophy of Science* 71 (2004), 805-819

James Woodward: Invariance, Modularity, and all That: Cartwright on Causation. Workshop-Paper Universität Konstanz 2002. www.uni-konstanz.de/ppm/workshop2/Abstracts.htm