Denise
D. Cummins

Gutes Denken

Wie Experten Entscheidungen fällen

SACHBUCH

Springer Spektrum

Gutes Denken

Die sieben Schlüsselkonzepte des Denkens

Wenn Sie dieses Buch gelesen haben, werden Sie in zweierlei Hinsicht weiser sein. Sie werden wissen, wie die besten und klügsten Denker entscheiden, argumentieren, Probleme lösen und richtig von falsch unterscheiden. Aber Ihnen wird auch bewusst sein, dass es durchaus nicht immer schlecht ist, wenn man diese Standards *nicht* erfüllt.

Denise D. Cummins stellt Ihnen die sieben entscheidenden Denkkonzepte vor, die die Welt verändert haben:

1. Denken lässt sich automatisieren, daher können wir Maschinen bauen, die denken.
2. Um Probleme zu lösen, sollten Sie immer Wege suchen, die den Abstand zwischen Ihrer aktuellen Situation und Ihrer Zielsituation verringern. Einsicht ist quasi eine implizite Suche.
3. Einige Gedanken führen zu weitergehenden Überlegungen, andere tun das nicht, und es gibt Regeln, mit denen Sie feststellen können, welche zur ersten Gruppe gehören und welche zur zweiten.
4. Um herauszufinden, was wahr ist, sollten Sie am besten zuerst herausfinden, was falsch ist.
5. Um zu entscheiden, welche Ursache etwas hat, ist es nötig, Alternativen zu bedenken.
6. Sie werden nicht immer bekommen, was Sie möchten, aber Sie können herausfinden, was Ihnen am ehesten dazu verhelfen wird.
7. Das Spiel ändert sich, wenn Sie es nicht allein spielen.

[Cummins] diskutiert, wie Ökonomen, Philosophen und andere Fachleute definiert haben, was eine Entscheidung rational oder ein Urteil moralisch macht. Sie legt die sieben Grundsätze des kritischen Denkens dar und erkundet die Taktiken, mit denen sich fehlerhafte Logik korrigieren lässt. **Scientific American**

Denise Dellarosa Cummins ist emeritierte Professorin für Psychologie und Philosophie an der University of Illinois in Urbana-Champaign. Zu ihren Forschungsschwerpunkten zählen Denken und Entscheidungsfindung unter evolutionären, vergleichenden und entwicklungspsychologischen Aspekten. Sie hat zahlreiche Fachartikel und ein populärwissenschaftliches Buch (*The Other Side of Psychology*) verfasst, ist Mitherausgeberin eines zweibändigen Werkes zur Kognitionswissenschaft und führt den Blog „Good Thinking" auf Psychology Today. Cummins hat an der Yale University, der University of California und am Max-Planck-Institut für Bildungsforschung in Berlin geforscht und gelehrt. Sie lebt mit Mann und zwei Töchtern in Champaign.

Denise D. Cummins

Gutes Denken

Wie Experten Entscheidungen fällen

Aus dem Englischen übersetzt von Regina Schneider

Springer Spektrum

Denise D. Cummins
Department of Psychology
202 Coble Hall
University of Illinois
Champaign, USA

Aus dem Englischen übersetzt von Regina Schneider.
Übersetzung der amerikanischen Ausgabe: Good Thinking – Seven Powerful Ideas That Influence the Way We Think von Denise D. Cummins, erschienen bei Cambridge University Press 2012,

ISBN 978-3-642-41809-9 ISBN 978-3-642-41810-5 (eBook)
DOI 10.1007/978-3-642-41810-5

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Springer Spektrum

Planung und Lektorat: Frank Wigger, Martina Mechler
Redaktion: Birgit Jarosch
Einbandabbildung: ©fotolia
Einbandentwurf: deblik Berlin

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Spektrum ist eine Marke von Springer DE. Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media
www.springer-spektrum.de

Inhalt

1
Einführung

Nachdem ich zwei Jahrzehnte lang kluge und wissbegierige Universitätsstudenten unterrichtet hatte, kam ich zu einem erschreckenden Befund: Trotz bester Bemühungen, den Studenten die Ideen und Erkenntnisse verständlich zu machen, die unsere heutige Denk- und Lebensweise entscheidend prägen, bleiben die meisten von ihnen doch sehr ihren einzelnen Disziplinen verhaftet. Studenten der naturwissenschaftlichen Studiengänge wissen alles über Hypothesentests, haben aber nicht die geringste Ahnung von Moraltheorie. Studenten der Philosophie und Rechtswissenschaften wissen alles über Beweisführung, haben aber nicht die geringste Ahnung von wissenschaftlichen Untersuchungsmethoden. Außerhalb der Wirtschaftshochschulen wissen nur herzlich wenige Studenten irgendetwas über Entscheidungstheorien, die den Aktienmarkt antreiben und wirtschaftspolitischen Strategien zugrunde liegen, die nicht zuletzt auch ihr jeweils ganz persönliches Leben beeinflussen (etwa in der Frage, wer einen Studentenkredit bekommt oder nicht). Und wer nicht gerade Psychologie studiert, weiß praktisch nichts darüber, wie die komplexen Verschaltungen der Nervenzellen in unserem Gehirn unser Denken, Fühlen und Handeln bestimmen. Nach dem Stu-

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_1,

dium starten diese klugen und gut ausgebildeten Menschen dann mit einem gepflegten Halbwissen ihre Karriere als politische Entscheidungsträger, Schriftsteller, Wissenschaftler, Juristen und Lehrer – wurschteln sich mit tiefen Wissenslöchern irgendwie durch, wo eigentlich elementare Grundkenntnisse vorhanden sein sollten.

Na und? Ist das denn so schlimm? Nun, im US-Bundesstaat Colorado gab es einmal einen Fall von Trunkenheit am Steuer, der trotz erdrückender Beweislage in einem Freispruch endete. „Mir völlig unverständlich", beklagte der Staatsanwalt. „Es war offenbar eine ausgemachte Sache. Der Blutalkoholspiegel des Burschen lag nachweislich über der gesetzlichen Promillegrenze, er konnte nicht mehr geradeaus auf einer Linie laufen und im Auto fand man leere Bierdosen mit seinen Fingerabdrücken darauf." Aber warum wurde er dann freigesprochen? „Nach dem Urteil sprach ich mit einem der Geschworenen", erzählte der Staatsanwalt weiter, „und der sagte mir, dass unter den Mitgliedern der Jury auch eine Astrologin war. Sie hatte ein Horoskop erstellt und argumentierte, dass der Angeklagte nach diesem Horoskop an jenem Tag keinesfalls im betrunkenen Zustand selbst gefahren sein konnte. Und so kam keine Mehrheit zusammen, die begründete Zweifel an der Unschuld des Angeklagten gehabt hätte."

Aber erzeugen Horoskope begründete Zweifel? „Nein, bestimmt nicht", würden die meisten von uns darauf antworten. Doch *warum* fällt unsere Antwort so entschieden aus, und warum stößt uns die Entscheidung der Jury so sehr vor den Kopf? Eine genaue Erklärung dafür zu finden, ist äußerst schwierig. Denn kaum einer würde wohl in ein Flugzeug steigen oder eine Brücke überqueren, ein Gesetz

befolgen oder sich mit Präsidentschaftswahlen befassen, wenn all diese alltäglichen Dinge auf derlei Begründungen und Beweismitteln basieren würden. Doch wir tun all das, weil wir voraussetzen, dass Flugzeuge, Brücken, Gesetze sowie unser Regierungssystem die Ergebnisse von genauesten Überlegungsprozessen, evidenzbasierten Bewertungen und Erkenntnissen aus vergangenen Fehlern sind. Wir glauben, dass die Vernunft das stählerne Garn ist, das unser Denken zu einem Bild verwebt, in dem Rechtsprechung gerecht und Wissenschaft exakt ist und in dem gesellschaftliche Einrichtungen robust und gegen Veränderungen gefeit sind. Kurzum, wir sind der tiefsten Überzeugung, dass Handlungen durch Vernunft gesteuert sein sollten.

Damit einher geht die grundlegende Überzeugung, dass Entscheidungen, die wir in einem emotional aufgewühlten Zustand treffen, schlechte Entscheidungen sind, während all jene, die wir im hellen und klaren Licht der Vernunft fällen, die besseren sind. Das erscheint uns schlicht logisch und selbstverständlich. Der folgende Wikipedia-Eintrag unter dem Begriff „Vernunft" fasst diesen weit verbreiteten Volksglauben folgendermaßen zusammen: „Der Begriff Vernunft bezeichnet in seiner modernen Verwendung die Fähigkeit des menschlichen Denkens, aus den im Verstand durch Beobachtung und Erfahrung erfassten Sachverhalten universelle Zusammenhänge in der Welt durch Schlussfolgerung herzustellen, deren Bedeutung zu erkennen, Regeln und Prinzipien aufzustellen und danach zu handeln." Und unter dem Begriff „Emotionalität" steht zu lesen, Gefühle seien bestimmt durch unklare Erkenntnisse und vernunftlose Gemütsbewegungen.

Diese Überzeugungen sind dermaßen fest verwurzelt, dass es uns oft überrascht zu erfahren, dass nicht jeder so denkt. Lee Harris beispielsweise schreibt in seinem Buch *The suicide of reason: radical Islam's threat to the west* (2007):

> „Der Westen hat ein Ethos des Individualismus, der Vernunft und der Toleranz entwickelt sowie ein wohldurchdachtes System, in welchem jeder Akteur, vom Einzelnen bis zum Gesamtstaat, Konflikte durch Worte zu lösen sucht. Das gesamte System basiert auf der Idee des Eigennutzes […] Unsere Verehrung der Vernunft macht uns zur leichten Beute für rücksichtslose, skrupellose und äußerst aggressive Prädatoren und trägt möglicherweise zu einem langsamen, kulturellen ‚Selbstmord' bei."

Für Philosophen wie Harris ist die Vernunft etwas, das uns schwach, unschlüssig, angreifbar und verwundbar macht. Die Vernunft verstrickt sich in unseren Worten und macht uns in unserem Handeln zögerlich und langsam. Es liegen zudem genügend Beweise vor, dass die menschliche Vernunft zerbrechlich und fehlbar ist. Wissenschaftler, die sich mit der menschlichen Vernunft und dem menschlichen Entscheidungsverhalten befassen, haben vielfach dokumentiert, mit welch dramatischer Häufigkeit wir fehlerhafte Entscheidungen treffen.

Aber nicht nur Wissenschaftler und politische Entscheidungsträger wissen um die Fehlbarkeit der menschlichen Vernunft, auf die sie sich in ihren Entscheidungen, die das Leben von Millionen von Menschen betreffen, nach wie vor stützen.

Die niederländische Politikwissenschaftlerin Ayaan Hirsi formuliert es so:

> „Aufklärer, die mit individueller Freiheit ebenso wie mit säkularer und begrenzter Regierungsgewalt beschäftigt waren, führten an, dass die menschliche Vernunft fehlbar ist. Sie erkannten, dass Vernunft mehr ist als bloß rationales Denken; es ist auch ein Prozess des Versuchs und des Irrtums, der Fähigkeit, aus den Fehlern der Vergangenheit zu lernen. Die Aufklärung kann nicht vollständig rezipiert werden, ohne ein klares Bewusstsein darüber, wie fragil der menschliche Verstand in der Tat ist. Genau deshalb sind Konzepte wie Zweifel und Reflexion so wesentlich für jegliche Formen vernunftbasierter Entscheidungen." („Blind faiths", *New York Times* 6.1.2008.)

Doch inwiefern sind die so überaus wichtigen Konzepte von „Zweifel und Reflexion" in unsere Entscheidungspraxis integriert? Es gibt Vernunftmodelle, die das westliche Denken bestimmen. Sie sind die „Juwelen" wissenschaftlicher Untersuchungsmethoden oder die Wissensbrücken, um einen Begriff zu verwenden, den der Philosoph Robert Cummins geprägt hat – Gedankengänge, die uns von dem, was wir bereits wissen, zu dem führen, was wir wissen wollen.

Sinn und Zweck dieses Buches ist, jede einzelne dieser Wissensbrücken in einfachen und verständlichen Worten darzulegen, damit der geneigte Leser für sich entscheiden kann, wie viel oder wie wenig wir auf die „Verehrung der Vernunft" geben sollten. Es geht um folgende Modelle:

1. Theorie der rationalen Entscheidung: wir wählen die Handlungsalternative aus, die unsere gewünschten Ziele größtmöglich realisiert
2. Spieltheorie: der Einzelne entscheidet nicht als Einziger
3. moralische Urteilsbildung: wie wir Richtig von Falsch unterscheiden
4. wissenschaftliche Argumentation, bestehend aus:
   - Hypothesentests: die Suche nach Wahrheit durch Bewertung von Evidenz
   - kausales Denken: Erklärungen, Vorhersagen und Verhinderung unerwünschter Ereignisse
5. Spiel der Logik: die Suche nach Wahrheit durch schlüssige Argumentation

Soweit die wichtigsten theoretischen Modelle, die den Entscheidungen, die wir in unserem Alltag, in der Rechtsprechung, in Politik, Wirtschaft und Wissenschaft treffen, zugrunde liegen. Bevor wir darüber entscheiden können, ob und inwieweit wir der Vernunft tatsächlich vertrauen können, müssen wir die Werkzeuge kennen, um zu verstehen, wie das „Spiel des rationalen Denkens" funktioniert, das von Experten so meisterlich gespielt wird.

Darüber hinaus gibt es zwei weitere Modelle, die eine Erörterung verdienen. Obgleich nicht so sehr formalisiert, wie die vier zuvor genannten, gehören sie zur Erforschung kognitiver Prozesse (zu denen Wahrnehmung, Erkennen, Urteilsvermögen, Problemlösen, logisches Schließen und Lernen zählen) bei menschlichen und anderen Individuen unbedingt dazu.

6. Problemlösung: die Suche nach Lösungen für unerwünschte Situationen
7. Analogieschlüsse: das Kernstück von Einsicht, Erkenntnis und Genius

Und schließlich: So rational und fehlerfrei diese Methoden scheinen mögen, sie werden nicht durch eine unfehlbare Hardware, einen unfehlbaren Erkenntnisapparat, realisiert. Vielmehr sind es menschliche Denker aus Fleisch und Blut, genauer gesagt, ihre neuronalen Schaltkreise, die diese Modelle erstellen. Um das gesamte „Vernunftpaket" vollauf erfassen zu können, müssen wir wissen, wie diese Schaltkreise unter verschiedenen Bedingungen funktionieren, damit wir am Ende zu komplexen Entscheidungen gelangen können. Aus diesem Grund werden in den folgenden Kapiteln wichtige Erkenntnisse aus den neuesten Forschungsbereichen der Neuro- und Kognitionswissenschaften, die für die einzelnen Denkmodelle relevant sind, ausführlich beschrieben.

Nach der Lektüre des Buches dürfte der Leser dann in der Lage sein, für sich selbst zu entscheiden, ob das vernunftgemäße Denken des Menschen tatsächlich so zerbrechlich oder so stark, so gefährlich oder so harmlos, so entbehrlich oder so wesentlich ist, wie es zu sein behauptet wird.

2

Spieltheorie

Der Einzelne entscheidet nicht als Einziger

John und Mary überlegen, wie sie ihren Freitagabend verbringen wollen. John würde lieber daheim bleiben und Videospiele spielen. Mary würde lieber ins Kino gehen. Beide aber wollen den Abend lieber zusammen als getrennt voneinander verbringen. Das Problem ist offensichtlich. Egal, wie sie sich entscheiden, einer von beiden zieht den Kürzeren: Wenn sie beide zu Hause bleiben und Videospiele spielen, ist John glücklich, aber Mary langweilt sich. Wenn sie zusammen ins Kino gehen, ist Mary glücklich, aber John hat das Nachsehen. Und wenn sie getrennte Wege gehen, sind beide unzufrieden.

Was also tun? Hier zu einer Entscheidung zu gelangen, ist sehr viel schwieriger, als es auf den ersten Blick scheinen mag, denn das Ergebnis wird nicht von einem Entscheider alleine bestimmt, sondern hängt für jeden einzeln davon ab, was der jeweils andere entscheidet – und beide wissen das. Schauen wir uns Mary und John ein bisschen näher an.

Es ist Montagnachmittag. Mary ist im Büro und tut alles, um einem lästigen Kollegen aus dem Weg zu gehen, der sich ständig mit ihr verabreden will, obwohl er weiß, dass sie verheiratet ist. Unweit ihrer Arbeitsstätte gibt es lediglich zwei Lokale, wo man in der Mittagspause etwas essen

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_2,

kann, Subway und Starbucks. Ginge Mary zu Subway und ihr Kollege ebenfalls, liefe sie ihm zwangsläufig über den Weg. Sie würde sich ärgern und er sich freuen. Dasselbe wäre der Fall, wenn sie beide zu Starbucks gingen. Ginge Mary zu Subway und ihr Kollege zu Starbucks, wäre sie erleichtert und er würde sich ärgern. Dasselbe wäre der Fall, wenn sie zu Starbucks ginge und er zu Subway. Wie das Ergebnis für jeden einzelnen ausfallen wird, hängt davon ab, was der jeweils andere tut – und beide wissen das.

In der Zwischenzeit sieht sich John mit einem ganz anderen Dilemma konfrontiert. Zusammen mit einem Kollegen hat er ein Gutachten gründlich verpfuscht, was seine Firma nun mit 100.000 Dollar teuer zu stehen kommen wird. Sein Chef ist außer sich vor Wut und erwägt, die beiden zur Rechenschaft zu ziehen und sich den Schaden von ihnen bezahlen zu lassen. Er zitiert beide getrennt voneinander zu sich und verlangt Auskunft darüber, wer genau den Schaden verursacht hat. Würden sie sich gegenseitig die Schuld anlasten, böte er sie beide mit jeweils 50.000 Dollar zur Kasse. Beschuldigte nur einer den anderen, würde er von dem Beschuldigten den vollen Betrag von 100.000 Dollar verlangen und der Kollege käme ungeschoren davon. Weigerten beide sich, den jeweils anderen zu beschuldigen, würde er von jedem 25.000 Dollar verlangen und die restlichen 50.000 abschreiben. John muss also entscheiden, ob er seinen Kollegen beschuldigen oder besser den Mund halten soll. Sein Kollege steht vor demselben Dilemma – und beide wissen das. Es ist eine Frage des Vertrauens: Wie das Ergebnis für jeden einzeln ausfallen wird, hängt davon ab, was der jeweils andere tut.

Wir alle kennen ähnliche Entscheidungssituationen aus dem eigenen Leben. Mathematiker nennen diese Art von Problemen (oder Entscheidungsdilemmas) „Spiele", und die damit verbundenen optimalen Entscheidungen können durch die sogenannte „Spieltheorie" bestimmt werden.

Grundlagen der Spieltheorie

Oskar Morgenstern und John von Neumann formulierten die grundlegenden Gedanken hinter der Spieltheorie in ihrem Buch *Spieltheorie und wirtschaftliches Verhalten* (Originaltitel von 1944: *The theory of games and economic behavior*). Zunächst einmal sind wir auf bestimmte Vermutungen angewiesen, die nach dem Satz von Bayes gebildet werden: Die Akteure oder *Spieler*, wie sie im mathematisch-formalen Sinne heißen, haben Präferenzen, die sie nach ihrem Nutzen (*Befriedigung*) ordnen, um dann in logischer Übereinstimmung danach zu handeln. (In der Spieltheorie steht der Begriff des Nutzens oft synonym für den Gewinn, den ein Spieler potenziell machen kann und der als Auszahlung bezeichnet wird. Daraus resultiert eine befriedigende Freude, die eine Person aus einem bestimmten Ergebnis, dem *Spielausgang*, zieht; Anm. d. Übers.)

Ein Spiel stellt eine Entscheidungssituation dar, in die mehr als ein Spieler eingebunden ist. Jeder Spieler versucht, seine „Auszahlungen" zu maximieren, doch wie hoch die Auszahlung für den einzelnen Spieler tatsächlich ausfällt, hängt davon ab, was der oder die anderen Spieler tut. Spiele werden definiert durch die Zahl der Spieler, die möglichen Handlungsoptionen, die dem einzelnen Spieler (Akteur)

zur Verfügung stehen, sowie die Zahl der möglichen Spielausgänge. *Konstantsummenspiele* beschreiben Situationen, bei denen die Höhe des Gesamtgewinns (die ausgezahlte Summe), den jeder einzelne Spieler erhalten kann, für alle möglichen Spielausgänge immer genau dieselbe ist. Man denke an TV-Sender, die um Zuschauer konkurrieren. Wenn es zehn Millionen Zuschauer gibt und drei Millionen davon sehen NBC, dann bedeutet das, dass die anderen Sender drei Millionen Zuschauer weniger haben. Wenn zwei Millionen davon auf ABC umschalten, gewinnt ABC zwei Millionen Zuschauer hinzu und NBC verliert zwei Millionen. Was der Gewinn des einen Spielers ist, ist der Verlust des anderen, und die Summe ist im Ergebnis immer dieselbe, egal welcher der beiden Sender die Zuschauer gewinnt und welcher sie verliert. In einem *Nullsummenspiel* (eine spezielle Form des Konstantsummenspiels) beträgt der Gesamtgewinn (die Auszahlungssumme) für alle Spieler immer genau Null. Wenn ich 1 Dollar gewinne, verlieren Sie 1 Dollar. Der Gewinn für mich ist also +1, für Sie ist er –1, und die Summe aus beidem beträgt Null. In einem *Nicht-Nullsummenspiel* kann die Summe aller Auszahlungen negativ oder positiv sein: Jeder Spieler kann einen Verlust erleiden oder jeder Spieler kann einen Gewinn einstreichen, aber die Summe aus Verlust und Gewinn ist für alle Spieler und für alle möglichen Spielausgänge immer dieselbe. Beispiel 1: Die Summe aller Auszahlungen beträgt 50 Dollar, egal wie das Spiel gespielt wird. Das bedeutet, dass jeder Spieler gewinnt: Wenn es nur zwei Spieler gibt, Sie und mich, und ich gewinne 30 Dollar, dann gewinnen Sie 20 Dollar. Beispiel 2: Die Summe aller Auszahlungen beträgt –50 Dollar, egal wie das Spiel gespielt wird. Das

bedeutet, dass jeder Spieler verliert: Wenn ich 30 Dollar verliere, verlieren Sie 20 Dollar.

Spiele können kooperativ oder nichtkooperativ sein. In *kooperativen Spielen* können die Spieler Koalitionen oder Allianzen bilden, um den erwarteten Nutzen zu maximieren. Beispiel Tennis: Ein Einzel ist ein nichtkooperatives Spiel – die Spieler spielen als Einzelspieler und konkurrieren darum, das Spiel zu gewinnen. Ein Doppel ist ein kooperatives Spiel – jeweils zwei Spieler spielen in einem Team und kooperieren, um das andere Team zu schlagen. Basketball, Football und Fußball sind allesamt Beispiele für kooperative Spiele (ein Freund von mir bezeichnet sie als „koalitionäre Ballbewegungsspiele"). Tennis-Einzel, Schachturniere sowie die meisten Videospiele sind nichtkooperative Spiele: Ein Einzelner konkurriert mit einem menschlichen Gegner oder einem Computer als Gegner, um das Spiel zu gewinnen.

In jeder Phase des Spiels tun die Spieler etwas – sie entscheiden sich für eine Aktion. Es kann viele verschiedene Spielausgänge geben, je nachdem, für welche Aktionen sich der einzelne Spieler entscheidet. Aktionen sind dabei als strategische Interaktionen zu verstehen. In einem Basketballspiel können die Spieler offensiv oder defensiv spielen. Sie können sich entscheiden, den Ball durch direkte Pässe über das Spielfeld zu bewegen, um ihn strategisch günstig zu positionieren. Dabei führen einige Aktionen zu besseren Ergebnissen für den einzelnen Spieler als andere. Die *beste Reaktion* eines Spielers ist die Strategie, die den höchstmöglichen Gewinn erbringt. Als Spieler oder Trainer der Mannschaft ist die beste Reaktion demnach die Strategie, die der Mannschaft am ehesten zum Sieg verhilft.

Hat das Spiel eine Phase erreicht, in der die einzelnen Spieler das Spielergebnis nicht mehr verbessern können, befindet sich das Spiel in einem *Gleichgewicht*. Jeder Spieler hat demnach eine Strategie verfolgt, die angesichts der Strategien der anderen Spieler das eigene Spielergebnis nicht verbessern konnte. Beispiel: Wenn ein Spieler oder ein Team ein Tennisspiel gewinnt, so hat das Spiel, wie es heißt, sein Gleichgewicht erreicht. Die Sieger können kein besseres Ergebnis mehr erzielen, sie haben das Spiel gewonnen. Und die Verlierer können kein besseres Ergebnis mehr erzielen, sie haben keine Punkte mehr zu gewinnen. Ein solches Gleichgewicht findet sich auch bei einem Remis im Schach, wo für keinen der beiden Spieler mehr ein Spielzug möglich ist, der seine Position verbessert. Das Spiel ist vorbei, aber keiner hat gewonnen.

Vergleichen wir damit eine Situation, die im Kinofilm *A Beautiful Mind – Genie und Wahnsinn* beschrieben wird: Einige junge Männer betreten eine Bar. Sie alle haben es auf die aufregendste Frau dort abgesehen, die jeder für sich erobern und mit nach Hause nehmen will. Alle buhlen um sie, aber nur einer kann gewinnen; all die anderen Frauen sind gekränkt und ziehen von dannen, sodass letztlich auch die übrigen Männer allein nach Hause gehen. Doch würden die Männer ihre Strategie ändern und sich nicht nur auf die begehrenswerteste Frau versteifen, sondern auch ein Auge auf die andere Frauen werfen, würden sie ihre Chancen auf eine nächtliche Eroberung, sprich einen Gewinn, um ein Vielfaches erhöhen. Mit anderen Worten: Die Männer könnten sehr viel besser abschneiden, indem sie ihre Strategien ändern – und das weiß jeder.

Bereits 1950 formalisierte der US-amerikanische Mathematiker diesen Gedanken für kooperative Spiele. In einem sogenannten *Nash-Gleichgewicht* sucht jeder Spieler für sich nach der besten Lösung oder nach der besten Antwort, wie es in der Spieltheorie heißt, und nimmt demnach zutreffend an, dass der andere Spieler dasselbe tut. Wenn jeder Spieler eine Strategie gewählt hat und kein Spieler mehr durch eine einseitige Änderung seiner Strategie profitieren kann, und solange alle anderen an ihrer Strategie festhalten, spricht man von einem *Nash-Gleichgewicht in reinen Strategien*. Um zu prüfen, ob ein solches Gleichgewicht vorliegt, muss man also lediglich überprüfen, ob einer der Spieler durch eine einseitige Änderung seiner Strategie seine Situation verbessern kann.

Spieltheorie und der Kampf der Geschlechter

Kehren wir zurück zu John und Mary und den Dilemmas, in denen sich befinden. Im ersten Dilemma würde John lieber Videospiele spielen als ins Kino zu gehen. Mary würde lieber ins Kino gehen als Videospiele zu spielen. Beide aber wollen den Abend lieber gemeinsam als getrennt voneinander verbringen. Dieses Problem wird in der Spieltheorie „Kampf der Geschlechter" genannt und hat eine sehr interessante Eigenschaft: Es gibt gleich zwei *Nash-Gleichgewichte in reinen Strategien*.

Der Kampf der Geschlechter ist nach obiger Beschreibung ein *simultanes Spiel*. Das heißt, die Spieler treffen ihre

Tab. 2.1 Kampf der Geschlechter in der Normalform.

		Mary	
		Kino	**Videospiele**
John	**Kino**	3,2	0,0
	Videospiele	0,0	2,3

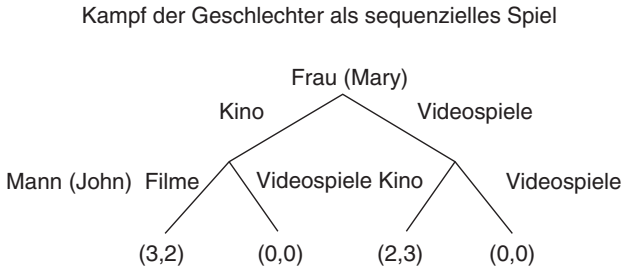
Entscheidung simultan und wählen ihre Lieblingsalternative zur selben Zeit, ohne zu wissen, wie sich der andere entschieden hat. Simultane Spiele werden oft in Form einer Matrix dargestellt, die Spielzüge und Auszahlungen (bei vollständiger Information) eines jeden Spielers beschreibt. Man spricht hier von der *Normalform* des Spiels, die für jede mögliche Strategiekombination das zugehörige Auszahlungsprofil angibt. Tabelle 2.1, dargestellt in Normalform, zeigt den Kampf der Geschlechter, den John und Mary auszufechten haben.

Marys erste Wahl heißt Kino, Johns erste Wahl heißt Videospiele – und beide wissen das. Was wäre, wenn beide sich auf ihre jeweils erste Wahl festlegten? Entschiede sich Mary für ihre erste Wahl (Kino), wüsste John, dass es für ihn die beste Wahl wäre, seine Strategie zu ändern und sich ebenfalls dafür zu entscheiden, ins Kino zu gehen. Entschiede sich John für seine erste Wahl (Videospiele), wüsste Mary, dass es für sie die beste Wahl wäre, ihre Strategie zu ändern und sich ebenfalls dafür zu entscheiden, den Abend zu Hause bei Videospielen zu verbringen. Es gibt hier also zwei Nash-Gleichgewichte in reinen Strategien: Kino-Kino und Videospiele-Videospiele. Gibt es eine Lösung, die aus dieser Sackgasse führt?

Ja. Die eine Möglichkeit besteht darin, dass John und Mary jeweils abwechselnd zum Zug kommen – das heißt,

einmal „gewinnt“ Johns erste Wahl, Mary „gewinnt“ beim nächsten Mal und so weiter. Der Kampf der Geschlechter wird so zu einem *sequenziellen Spiel*. In einem sequenziellen Spiel entscheiden die Spieler ihre Spielzüge abwechselnd und wissen jeweils, welche Züge bereits vorgenommen wurden. Nehmen wir einmal an, John und Mary schreiben jedes Mal auf, ob sie lieber ins Kino gehen oder Videospiele spielen, sodass sie beide immer wissen, wo in diesem Spiel sie gerade stehen. Wenn jeder Spieler beobachten kann, welchen Spielzug der jeweils andere vor ihm macht, ist es ein Spiel der *vollständigen Information*. Nehmen wir nun stattdessen an, John und Mary schreiben ihre jeweiligen Entscheidungen nicht auf, und nehmen wir weiterhin an, dass Mary sich die Spielzüge sehr viel besser merken kann als John. Wenn einige (aber nicht alle) Spieler also Informationen über vorausgegangene Spielzüge haben, ist es ein Spiel der *unvollständigen Information*. Sequenzielle Spiele werden anhand eines sogenannten *Spielbaums* beschrieben, der für jeden Spielzug und jede mögliche Antwort die zugehörigen Auszahlungsprofile (und Informationen) anzeigt. Diese Art der Darstellung wird als *Extensivform* bezeichnet und beinhaltet eine vollständige Beschreibung des Spiels, einschließlich der Reihenfolge möglicher Spielzüge, Auszahlungen und Informationen, die für jeden einzelnen Spieler für jeden einzelnen Zug verfügbar sind. Abbildung 2.1 zeigt den Kampf der Geschlechter im Fall von John und Mary in der Extensivform.

John und Mary könnten stattdessen beschließen, der Sackgasse zu entkommen, indem sie eine Münze werfen. Was wäre dann? Bringen die beiden ein Zufallselement in



Der Kampf der Geschlechter als sequenzielles Spiel, in dem die Spieler abwechselnd über ihre Spielzüge entscheiden und die vorangegangenen Spielzüge kennen.

Abb. 2.1 Kampf der Geschlechter in der Extensivform.

das Spiel ein, handelt es sich formal nicht um eine reine Strategie, sondern um ein *Nash-Gleichgewicht in gemischten Strategien*, und die erste Wahl eines Spielers ist damit auf die Wahrscheinlichkeiten reduziert, die mit dem eingebrachten Zufallselement einhergehen. Da Mary und John sich nun entscheiden, in jedem gegebenen Spiel eine Münze zu werfen, haben sie jeweils eine 50%ige Chance, dass die Entscheidung auf ihre jeweils erste Wahl fällt. Sie könnten aber auch Schere-Stein-Papier spielen und ihre jeweilige Gewinnchance damit auf 2:3 zu erhöhen. In jeder Spielrunde liegt die Gewinnchance dann bei 1:3. Und wenn sie nun umschwenken und lieber Streichhölzer ziehen wollen, liegt die ihre jeweilige Gewinnchance bei 1:Gesamtzahl der Streichhölzer.

Und an diesem Punkt gelang dem genialen Mathematiker John Nash, dem „Beautiful Mind", wie er in Anlehnung an den preisgekrönten Spielfilm, der seine Lebensgeschichte skizziert, auch genannt wird, ein geradezu brillan-

ter Beweis: Wenn es eine endliche Zahl von Spielern und eine endliche Zahl von Strategien in einem Spiel gibt, muss es wenigstens ein Nash-Gleichgewicht geben – entweder nach dem Konzept der reinen Strategie (man wählt eine Strategie und zieht sie durch) oder nach dem Konzept der gemischten Strategie (man führt ein Zufallselement ein). Seine Beiträge zur Spieltheorie brachten ihm 1994 den Nobelpreis für Wirtschaftswissenschaften ein (zusammen mit John Harsanyi und Reinhard Selten).

Spieltheorie und Mary, die versucht, ihrem lästigen Kollegen aus dem Weg zu gehen

Kommen wir nun zum zweiten Dilemma im Fall von John und Mary, respektive zu Mary, die versucht, einen Bogen um ihren Kollegen zu machen. Dieses Spiel trägt den Namen *Matching Pennies* und weist ebenfalls eine sehr interessante Eigenschaft auf: Im Matching-Pennies-Spiel gibt es kein Nash-Gleichgewicht.

Das Spiel heißt Matching Pennies, weil es dieselbe Struktur hat wie das Spiel, das wir aus Kindertagen kennen: Jeder von zwei Spielern hat eine Münze in der Hand und entscheidet sich vor dem Wurf für Kopf oder Zahl. Der eine Spieler gewinnt, wenn beide Münzen dasselbe zeigen (also zweimal Kopf oder zweimal Zahl), und der andere gewinnt, wenn sie unterschiedlich fallen (also einmal Kopf und einmal Zahl). Sagen wir mal, Sie gewinnen, wenn beide Münzen dasselbe zeigen, und Ihr Freund gewinnt, wenn

Tab. 2.2 Matching-Pennies-Spiel in der Normalform.

		Spieler 1	
		Kopf	**Zahl**
Spieler 2	**Zahl**	+1, –1	–1, +1
	Kopf	–1, +1	+1, –1

sie unterschiedlich fallen. Wenn Sie beide auf Kopf setzen, werden Sie immer gewinnen und Ihr Freund wird immer verlieren. Ihr Freund hat demnach einen Anreiz, seine Strategie zu ändern und auf Zahl zu setzen. Aber das bedeutet, dass nun Sie derjenige sind, der immer verliert und einen Anreiz hat, fortan auf Zahl zu setzen. Jetzt ist es jedoch Ihr Freund, der immer verliert und einen Anreiz hat, sein Verhalten zu ändern, und so weiter und so fort. Es gibt für dieses Spiel also kein Nash-Gleichgewicht in reinen Strategien. Tabelle 2.2 stellt dieses Spiel in Normalform dar.

Was tun Sie also? Wahrscheinlich das, was jedes Kind in einer solchen Situation macht: Sie wechseln zufällig zwischen Kopf und Zahl hin und her. Damit liegt die Wahrscheinlichkeit zu gewinnen für jeden Spieler bei 50 %.

Wenn man also ein Zufallselement in ein Spiel einbringt, für das kein Gleichgewicht nach reinen Strategien existiert, ergibt sich, wie Nash bewiesen hat, ein Gleichgewicht in gemischten Strategien. Beide Spieler haben nun eine 50%ige Gewinnchance und könnten genauso gut an ihrer jeweiligen Strategie festhalten. Und Mary? Sie könnte es genauso machen und eine Münze werfen, um zu entscheiden, ob sie in der Mittagspause zu Subway oder Starbucks geht.

Spieltheorie und die Frage nach der kommunikativen Glaubwürdigkeit

Zurück zu John und damit zum dritten Dilemma: Was soll John am besten tun? Soll er dichthalten und mit seinem Gegenspieler, sprich seinem Kollegen, kooperieren (*Kooperation*), oder soll er sich aus der Affäre ziehen und ihn verraten (*Defektion*)? Dieses Spiel ist bekannt als das sogenannte Gefangenendilemma, denn es skizziert ein Szenario, in dem die Indizienbeweise nicht ausreichen, um beide Beschuldigten eines gemeinsam begangenen Verbrechens zu überführen. Bringt man aber einen von beiden dazu, gegen den anderen auszusagen, kann dieser für schuldig befunden werden. Ein Anwalt würde versuchen, einen Handel anzubieten. John käme beispielsweise ungeschoren davon, wenn er gegen seinen Kollegen aussagt. Auch dieses Spiel hat eine interessante Eigenschaft: Es bildet stets ein *Nash-Gleichgewicht in dominanten Strategien*.

Im Gefangenendilemma hat jeder Spieler eine dominante Strategie – das bedeutet, dass die beste Antwort des einzelnen Spielers nicht abhängig ist von der Strategie des anderen Spielers. Egal, was der andere Spieler tut, es gibt eine Strategie, die für jeden einzelnen am besten funktioniert. Wenn beide Spieler über dominante Strategien verfügen, die koinzidieren, befindet sich das Spiel in einem Gleichgewicht dominanter Strategien, einer Sonderform des Nash-Gleichgewichts (denn jedes Gleichgewicht dominanter Strategien macht ebenso gleichzeitig ein Nash-Gleichgewicht sichtbar; Anm. d. Übers.). In einem einmalig gespielten Gefangenendilemma gibt es nur zwei mögliche

Tab. 2.3 Gefangenendilemma in der Normalform.

		Kollege	
		Defektion	**Kooperation**
John	**Defektion**	P = –50.000 $, –50.000 $ Strafe für gegenseitige Defektion (P = *punishment*)	T = 0 $, –100.000 $ Versuchung zu defektieren (T = *temptation*)
	Kooperation	S = –100.000 $, 0 $ Auszahlung nach Kooperation mit einem Defektor (S = *sucker's payoff*)	R = –25.000 $, –25.000 $ Belohnung für gegenseitige Kooperation (*reward*)

Antworten (Kooperation oder Defektion) und daher nur zwei mögliche Strategien (Kooperation oder Defektion). Das Dilemma von John und seinem Kollegen lässt sich in Normalform darstellen (Tab. 2.3).

Sehen wir uns die erste Zeile in der Tab. 2.3 an. Wenn sich John und sein Kollege gegenseitig beschuldigen, bezahlen sie jeder 50.000 Dollar. Wenn nur John seinen Kollegen beschuldigt, dieser aber schweigt, bezahlt John nichts und der Kollege die volle Summe von 100.000 Dollar. Sehen wir uns nun die untere Zeile an. Wenn stattdessen John schweigt, sein Kollege ihn aber beschuldigt, bezahlt John die volle Summe von 100.000 Dollar, und wenn sie beide schweigen, zahlt jeder bloß 25.000 Dollar.

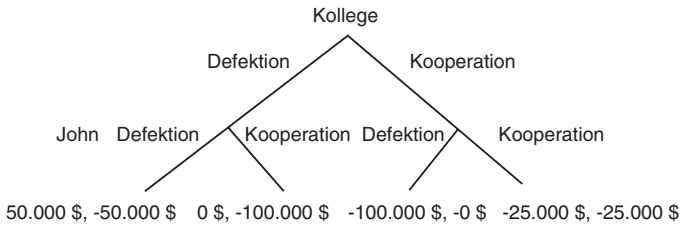
Beide sind sich der Situation bewusst und haben keinerlei Einfluss darauf, wie der jeweils andere handelt. Jeder geht davon aus, dass der andere das macht, was ihm den größten Vorteil verschafft. Demnach gibt es für jeden Spieler in einem einmalig gespielten Gefangenendilemma nur eine dominante Strategie, und die heißt Defektion.

„Aber halt", könnten Sie denken. Die zwei hätten im Vorfeld absprechen können, dass sie beide schweigen, womit die beste (oder fairste) Entscheidung für John darin bestünde, einfach den Mund zu halten. Aber wer sagt ihm, dass sich sein Kollege auch an die Absprache hält und nicht umfällt, wenn es Spitz auf Knopf steht? Die beste Entscheidung ist also immer noch die, sich selbst vor Strafe zu schützen. Solch eine Strategie zielt im Kern darauf ab, den eigenen Vorteil zu vergrößern, und geht davon aus, dass der andere ebenfalls nach dieser für ihn besten Strategie agieren wird.

Aber was wäre, wenn John und sein Kollege beste Freunde wären und ihre Zusammenarbeit stets von Kooperation geprägt gewesen wäre? Oder was wäre, wenn John und sein Kollege Konkurrenten wären und ihre Zusammenarbeit stets von heftigen Rivalitäten geprägt gewesen wäre? Würde die gemeinsame Vorgeschichte die Natur des Spiels ändern? Welche Strategie sollte John als die für ihn beste wählen – Kooperation oder Defektion? Oder gäbe es gar keinen Unterschied?

Doch, einen sehr großen sogar. Und zwar dann, wenn wir für Johns Situation ein wiederholtes Gefangenendilemma betrachten, denn in einer fortlaufenden Beziehung mit seinem Kollegen kann John die aus seiner Sicht richtige Strategiekombination vom Verhalten seines Kollegen in der Vergangenheit abhängig machen. In einem wiederholten Spiel ist jedem Spieler die jeweils vorherige Entscheidung des anderen bekannt. Das Gefangenendilemma lässt sich hier in Extensivform darstellen (Abb. 2.2).

Robert Axelrod, Professor der Politikwissenschaften an der University of Michigan, untersuchte 1980 das Gefange-



Das Gefangenendilemma als sequenzielles Spiel, in dem die Spieler ihre Spelzüge abwechselnd entscheiden und jeweils wissen, welche Züge bereits vorgenommen wurden.

Abb. 2.2 Wiederholtes Gefangenendilemma in der Extensivform.

nendilemma unter dem Aspekt der Kooperation und führte dazu Turniere mit zahlreichen Spielen durch, in denen Computerprogramme über mehrere Runden wiederholt gegeneinander antraten. Das Programm mit dem besten Ergebnis sollte am Ende gewinnen. Jedes Programm verfolgte eine Spielstrategie, die, bezogen auf die vorherigen Spielzüge (von Spieler und Gegenspieler), entweder auf Kooperation oder Defektion setzte. Einige der in den Turnieren eingesetzten Strategien waren die folgenden:

- Defektiere immer: Diese Strategie defektiert in jeder Runde. Es handelt sich um ein einfaches spieltheoretisches Modell und ist die sicherste Strategie, da sie dem einzelnen Spieler keinen Vorteil bringt. Jedoch versäumt sie die Chance, größere Gewinne zu erzielen, indem sie mit dem Gegenspieler kooperiert, der zu kooperieren bereit ist.
- Kooperiere immer: Diese Strategie funktioniert sehr gut, wenn sie beidseits gespielt wird. Entscheidet sich der

Gegenspieler allerdings zu defektieren, geht die Strategie nicht auf.
- Zufall: Diese Strategie kooperiert in 50 % über den gesamten Spielverlauf.
- Tit for Tat (Wie du mir, so ich dir): In dieser Strategie beginnt ein Spieler mit einem kooperativen Spielzug und wählt für den nächsten eigenen Zug dann die Strategie des letzten Zugs des Gegenspielers.

Die ersten drei Strategien sind im Vorhinein festgelegt, weshalb kein Spieler einen Vorteil erlangen kann, wenn er die vorangegangenen Spielzüge seines Gegenspielers kennt und seine Strategie durchschaut. Das heißt, die Spieler erfahren im Verlauf des Spiels nichts über ihren jeweiligen Gegenspieler. Die vierte Strategie, Tit for Tat, allerdings modifiziert das Verhalten der Spieler, da jeder die jeweils letzte Entscheidung seines Gegenspielers in der Vorrunde – und nur in dieser – kennt.

Wie die 1981 veröffentlichten Ergebnisse des Turniers zeigten (Axelrod & Hamilton 1981), war die insgesamt erfolgreichste Strategie Tit for Tat. Diese Strategie streicht sämtliche Vorteile der Kooperation ein, wenn sie gegen einen freundlichen (kooperativen) Gegenspieler antritt, riskiert zugleich aber nicht, übervorteilt zu werden, wenn der (unfreundliche) Gegenspieler defektiert. Es sei darauf hingewiesen, dass es sich bei der Tit-for-tat-Strategie um eine unbewusste Anwendung von intelligenten Heuristiken handelt, wie sie von Gigerenzer und seinen Kollegen untersucht wurden; sie ignoriert einen Teil der verfügbaren Information, erfordert praktisch keine Berechnungen und ist dennoch höchst erfolgreich. (Mehr dazu in Kapitel 3 über die intelligenten Heuristiken von Gigerenzer.)

Es gab noch weitere interessante Ergebnisse. Wenn zwei „freundliche" Tit-for-Tat-Spieler aufeinandertreffen, kooperieren sie immer. Der Grund: Das erste Spiel beginnt immer mit einem kooperativen Spielzug, auf den nach dem Prinzip Tit for Tat ebenfalls ein kooperativer Spielzug folgt.

Beantwortet ein Spieler diesen Zug beim nächsten Zusammentreffen entgegengesetzt und defektiert, so kooperieren die Tit-for-Tat-Spieler nur im ersten Spielzug und verlegen sich dann auf eine fortdauernde Defektion. Und wird auf beiden Seiten jeweils nach der Zufallsstrategie defektiert oder kooperiert, so richten die Spieler ihre Zugfolgen nach denen ihres Gegenspielers. Aus diesem Grund kann die Tit-for-tat-Strategie nicht als die unbedingt „beste" Strategie bezeichnet werden, war aber stark genug, das Turnier zu gewinnen.

Was aber soll nun John tun? Wenn die Vorgeschichte erkennen lässt, dass sein Kollege immer kooperiert, dann sollte John kooperieren. Wenn er hingegen immer defektiert, dann sollte auch John defektieren. Und wenn sein Kollege unberechenbar ist, sollte er ebenfalls defektieren. Was also lernen wir daraus? Ein Spieler sollte versuchen, die Strategie seines Gegenspielers zu ergründen (oder zu erahnen), um dann eine Strategie zu wählen, die für die Situation am besten geeignet ist.

Experimentelle Ökonomie: Wie verhalten wir uns tatsächlich?

Was passiert tatsächlich, wenn das Gefangenendilemma von realen Menschen in realen Situationen gespielt wird? In Studien der experimentellen Wirtschaftsforschung spie-

len Menschen gegen andere Menschen, und echtes Geld wechselt den Besitzer. Am Ende einer Studie geht jeder mit seinem erspielten Gewinn nach Hause.

Wir erinnern uns: Eine erfolgreiche Strategie hängt nach spieltheoretischer Sichtweise davon ab, dass ein rationaler Akteur erstens gemäß dem eigenen Interesse agiert, dass er zweitens das eigene Interesse nach dem maximalen Gewinn für sich selbst definiert und dass er drittens in Spielen, für die es eine solche gibt, eine dominante Strategie verfolgt und viertens in Spielen, für die es ein solches gibt, nach einem Nash-Gleichgewicht strebt.

Wie wir eben gesehen haben, besteht die rationale Entscheidung im einmaligen Gefangenendilemma in der Defektion. Defektieren Sie, würde ein Wirtschaftswissenschaftler Sie als einen rationalen Spieler bezeichnen, da Sie die Strategie spielen, die Ihnen am wahrscheinlichsten maximale Gewinne beschert. Umso überraschender mag es für Sie sein, wenn ich Ihnen sage, dass die Spieler in einem einmalig gespielten Gefangenendilemma über rund die Hälfte der Spielzeit kooperieren (Camerer 2003). Fragt man sie, warum sie dies tun, so nennen sie nicht etwa Altruismus als Grund. Stattdessen geben sie an, sie würden erwarten, dass der andere ebenfalls kooperiert nach dem Motto: „So wie ich es mache, wird der andere es auch machen." Ein solches Verhalten geht eher konform mit dem Prinzip der Reziprozität („Ich helfe dir, wenn du dafür mir hilfst"), ein Altruismus, der auf Gegenseitigkeit beruht, als mit einem unilateralen Altruismus („Dann nimm dir halt alles Geld, ist mir egal"). Wir können also sagen, dass die Spieler mit einer Tendenz zur Kooperation an diese Spiele herangehen, eine Tendenz, die für Wirtschaftswissenschaftler „irratio-

nal“ ist (Kelley & Stahelski 1970). Tatsächlich scheinen wir bei diesen Interaktionsformen die *Norm* der Kooperation zu befolgen.

Soziale Normen stellen Verhaltensstandards dar, die definiert sind durch die weitgehend gemeinsam vertretenen Erwartungen einer Gruppe an ihre einzelnen Mitglieder im Hinblick auf die richtigen Verhaltensweisen in einer gegebenen Situation. Sie sind üblicherweise als Soll-Regeln formuliert, die vorgeben, welches konkrete Verhalten in einer bestimmten Situation *erlaubt*, *vorgeschrieben* oder *verboten* ist. Meist entstehen Forderungen nach sozialen Normen dann, wenn das Handeln des Einzelnen für andere Menschen positive oder negative Nebeneffekte hat. Wir richten uns freiwillig nach sozialen Normen, wenn diese den eigenen Zielen (Eigeninteressen) entsprechen. Und wir richten uns unfreiwillig nach sozialen Normen, wenn das Interesse der Gruppe mit unserem eigenen in Konflikt gerät sowie auch dann, wenn andere die Macht haben, Normen zwangsweise durchzusetzen und mithin nichtkooperierende Mitglieder der Gruppe zu bestrafen.

Sehen wir uns einmal an, wie sich Probanden in Experimenten mit gemeinschaftlichen Gütern verhalten, in sogenannten Öffentliches-Gut-Spielen (Box 2.1). Für solche Gemeingüter werden Beiträge geleistet. Die Gemeingüter können von allen Mitgliedern einer Gemeinschaft genutzt werden, auch von denen, die selbst keinerlei Beitrag entrichten. Schön und gut, bis auf das Problem der Trittbrettfahrer, die von diesen Gütern nutznießen, ohne selbst etwas beizutragen. Jedes Mitglied hat einen Anreiz, als Trittbrettfahrer (oder Schmarotzer) von den Beiträgen anderer zu profitieren. Experimentell lassen sich solche Situationen als

Gefangenendilemma modellieren, wobei die Kooperation die konditionale Strategie ist (man kooperiert nur dann, wenn alle anderen auch kooperieren, und sonst nicht) und die Defektion die Trittbrettstrategie. Ebenso wie in den Experimenten zum Gefangenendilemma hängt das Ereignis der Defektion davon ab, ob diejenigen, die defektieren (die Trittbrettfahrer), bestraft werden oder nicht.

Box 2.1 So funktionieren Öffentliches-Gut-Spiele

Spielkomponenten

- Gruppen bestehen aus n Einzelspielern, wobei n größer ist als 2.
- Jeder Einzelspieler wird mit Geldeinheiten ausgestattet (E).
- Die Einzelspieler entscheiden, wie viel von E sie für sich selbst behalten und wie viel sie an ein Gruppenprojekt geben wollen.
- Der Experimentator multipliziert die Gesamtsumme, die für das Gruppenprojekt gegeben wurde, mit einer Zahl b, die größer ist als 1, aber kleiner als n.
- Die multiplizierte Summe der Beiträge der einzelnen Spielmitglieder ergibt den Gewinn aus dem Gruppenprojekt.
- Dieser Gewinn werden dann gleichmäßig unter den n Einzelspielern verteilt.

Ergebnisse

Wenn alle Einzelspieler ihre Geldeinheiten behalten, gewinnt jeder von ihnen E. Wenn alle ihre gesamten Geldeinheiten beisteuern, beträgt die Summe der Einlagen nE, und

- es ergibt sich ein Gewinn von $(b/n)nE = bE$
- der für jeden Einzelspieler größer ist als E.

Beispiel

$$E = 20,\ b = 2,\ n = 4;\ \text{Gewinn}: (b/n)nE = bE$$

Wenn jeder Spieler Null Geldeinheiten beisteuert, gewinnt jeder 20. Spendiert jeder Spieler seine gesamten Geldeinheiten, gewinnt jeder (2/4)4(20)=2(20)=40. Wenn jeder 5 Geldeinheiten beisteuert, gewinnt jeder (2/4)4(5)=10, plus 15, die jeder für sich behält=25.

Das Gefangenendilemma ist eine spezielle Variante des Öffentliches-Gut-Spiels mit n=2 und zwei verfügbaren Aktionen: nichts beisteuern (defektieren) oder alles beisteuern (kooperieren). Beide Spieler sind im Gefangenendilemma besser dran, wenn sie defektieren (weil b/n < 1), ganz gleich, was der Gegenspieler tut.

Eine solche Studie wurde von Fehr und Gächter (2000) beschrieben. Die Teilnehmer spielten mit Geldeinheiten, die sie in echtes Geld eintauschen konnten. In jedem Durchgang des Experiments bekamen sie die Möglichkeit, ihr Geld zu behalten oder aber einen Teil davon (oder alles) für ein gemeinschaftliches Projekt herzugeben. Für jede Geldeinheit, die sie behielten, bekamen sie eine weitere dazu. Für jede Geldeinheit, die sie dem Projekt beisteuerten, bekam jeder der Spieler – auch, die, die nichts beisteuerten – eine Viertel Geldeinheit dazu. Am Ende des Experiments wurden die Geldeinheiten eines jeden Spielers nach einem allgemein bekannten Wechselkurs in echtes Geld umgetauscht. Im ersten Durchgang bewegten sich die investierten Beiträge um die 50 % der Summe, die jedem zur Verfügung steht. Danach verlegten sich die Spieler zunehmend

auf das Trittbrettfahren. Die Spieler begannen, Geldeinheiten dazuzuverdienen, ohne von ihrem eigenen Geld auch nur eine Einheit beizusteuern. Je länger das Spiel andauerte, desto weniger steuerten sie bei, und bis zur zehnten Spielrunde waren die Beiträge gänzlich versiegt.

In der elften Runde dann bekamen die Spieler die Möglichkeit, die Trittbrettfahrer zu bestrafen, indem sie ihnen ein Bußgeld auferlegten, das in Geldeinheiten zu bezahlen war. Die Ergebnisse waren beachtlich: Die Kooperation schnellte prompt auf 60 %, kletterte danach kontinuierlich weiter und lag bis zur 20. Runde wieder bei 100 %. Wie die Ergebnisse zeigen, steuern wir weniger bis gar nichts bei, wenn Trittbrettfahren toleriert wird.

Genau wie in einem Gefangenendilemma mit zwei Akteuren zeigen die Spieler auch in Öffentliches-Gut-Spielen eine starke Neigung, andere zu bestrafen, um die Kooperation zu befördern – auch wenn sie dafür einige Kosten selbst tragen müssen. Eine Studie von Fehr und Fischbacher (2004a) war so angelegt, dass der Bestrafende wie auch der Bestrafte eine Geldeinheit zu bezahlen hatte, und zwar jedes Mal, wenn ein Trittbrettfahrer bestraft wurde. Wie sich herausstellte, waren die Spieler bereit, bis zu zwei Geldeinheiten zu bezahlen, um einen Trittbrettfahrer zu bestrafen. Unter Wirtschaftswissenschaftlern sorgte dieses Ergebnis durchaus für Verwunderung, denn es bedeutet, dass das Verlangen, Normverletzungen zu bestrafen, stärker ist als das, die eigenen Interessen zu verfolgen.

Die Androhung von Strafmaßnahmen hat überdies großen Einfluss auf das Verhalten der Menschen in sogenannten Vertrauensspielen. In dieser Art von Spiel können die

Akteure in die Rolle von Investmentbankern schlüpfen, die insbesondere Treuhandkonten verwalten. In einem einmaligen Vertrauensspiel geht es nun darum, dass der Investor einen Teil seines Geldes beim Treuhänder (Vertrauensnehmer) investiert. Der Investor kann dem Treuhänder so viel oder so wenig Geld anvertrauen wie er möchte, doch wird der transferierte Geldbetrag vom Experimentator verdreifacht. Diese Verdreifachung simuliert den Investitionsgewinn. Gibt der Investor also 1 Dollar an den Treuhänder, bekommt dieser tatsächlich 3 Dollar. Dem Treuhänder steht es sodann frei zu entscheiden, wie viel von dem gewonnenen Geld (oder auch gar nichts) er an den Investor wieder zurückzahlen will.

Würden die Treuhänder nach rein eigennützigen Überlegungen handeln, behielten sie das gesamte Geld für sich. Aber das tun sie nicht. Üblicherweise transferieren die Investoren ungefähr die Hälfte ihres Geldes an den Treuhänder, und die Treuhänder geben etwas weniger als die Hälfte wieder an die Investoren zurück. Doch was passiert, wenn die Investoren die Treuhänder sanktionieren könnten, wenn die Rendite geringer ausfällt als eine zuvor definierte Rendite. Fehr und Rockenbach (2003) untersuchten diese Bedingungen im Experiment und fanden heraus, dass die Treuhänder 50 % mehr Geld zurückbezahlten, wenn die Rendite es ihnen erlaubte, selbst Geld damit zu verdienen, aber 67 % weniger Geld zurückbezahlten, wenn sie ihnen Kosten verursachte. Man spricht hier bisweilen von der *Norm der gerechten Sanktion* (nach dem Motto: „Okay, ich denke, meine Gier verdient eine kleine Strafe, aber so eine große Strafe ist ungerecht!“). Fehr und Rockenbach kamen

zu dem Schluss, dass das altruistische Verhalten zurückgeht, wenn eine Strafe zu hart oder zu ungerecht ausfällt.

Wir halten also fest: Das menschliche Verhalten in wiederholten Spielen gibt den Wirtschaftsforschern einige Rätsel auf, da wir Kooperation offenbar großzügiger belohnen und Defektion härter bestrafen, als es durch spieltheoretische Analysen prognostiziert wird (Weg & Smith 1993). Wie die Studien zeigen, wichen die Ergebnisse besonders dann sehr stark von den spieltheoretischen Prognosen ab, wenn die Teilnehmer in Gefangenendilemma-Spielen, Öffentliches-Gut-Spielen und Vertrauensspielen die Möglichkeit hatten, Trittbrettfahrer zu entlarven und zu bestrafen. Insofern neigen wir also nicht bloß zur Kooperation, sondern wir erwarten sie auch von anderen und rächen uns empfindlich, wenn die anderen sich nicht kooperativ verhalten.

Für eine Wirtschaftstheorie, die auf Eigeninteresse basiert, bedeutet das Folgendes: Wer im Gefangenendilemma defektiert, wird bestraft, obgleich er sich so verhält, wie es der Spieltheorie zufolge die optimale Strategie ist. Nicht unmittelbar beteiligte Spieler gehen sogar so weit, selbst Kosten in Kauf zu nehmen, um die Auszahlung an einen anderen zu reduzieren und ihn dadurch zu bestrafen, wenn er in einem beobachteten Gefangenendilemma defektiert, vor allem dann, wenn der andere defektiert und sein Gegenspieler nicht (Fehr & Fischbacher 2004b). Kurzum: Das menschliche Verhalten ist nicht nur durch wirtschaftliches Eigeninteresse in Gefangenendilemma-Spielen motiviert, sondern auch durch Normen der Fairness und einem erwarteten Altruismus, der auf Gegenseitigkeit beruht (Reziprozität).

Unterschiede in Macht und Status beeinflussen, wie fair wir andere behandeln

Die experimentelle Wirtschaftsforschung kennt zwei weitere Spiele, in denen das menschliche Verhalten ebenfalls in Widerspruch zur Theorie steht – das Diktatorspiel und das Ultimatumspiel. In einem einmaligen Diktatorspiel gibt es zwei Akteure (Spieler), von denen einer, der Diktator, die uneingeschränkte Entscheidungsgewalt bekommt und einseitig bestimmen kann, wie ein bestimmter Geldbetrag zwischen den beiden Spielern aufgeteilt wird. Der zweite Akteur kann darauf keinerlei Einfluss nehmen. Nach der Spieltheorie müsste ein rationaler Diktator den gesamten Geldbetrag für sich behalten. Doch das tut er nicht. Wie experimentelle Studien zeigen, bieten die Spieler, die die Rolle des Diktators haben, ihrem Gegenspieler üblicherweise einen Betrag zwischen 15 und 35 % an (Camerer 2003). Das menschliche Verhalten wird hier von einem reinen Altruismus beeinflusst. Es handelt sich um ein einmaliges Zweipersonenspiel, in denen die beiden Akteure (wie auch der Experimentator) die Identität des jeweils anderen nicht kennen. Trotzdem geben die Spieler Geld an diese fremde Person, das sie ebenso gut und ohne Furcht vor Konsequenzen in die eigene Tasche hätten stecken können.

In einem einmaligen Ultimatumspiel kommt einem von zwei Spielern ebenfalls die Rolle zu, einen bestimmten Geldbetrag beliebig auf sich und den zweiten Spieler aufzuteilen. Doch es gibt einen entscheidenden Knackpunkt: Der zweite Spieler kann mitbestimmen. Der Spieler, der das

Geld verteilt, hat die Rolle des Vorschlagenden. Er macht ein Angebot. Der andere hat die Rolle des Empfängers. Er kann das Angebot annehmen oder ablehnen. Nimmt er das Angebot an, wird das Geld wie vorgeschlagen aufgeteilt und beide trennen sich. Schlägt er das Angebot aus, erfolgt keine Auszahlung und beide gehen leer aus.

Gehen wir davon aus, dass Individuen sich ausschließlich egoistisch und rational verhalten, müsste ein rationaler Vorschlag bei etwas mehr als Nichts liegen und ein rationaler Empfänger müsste auf jedes ihm vorgeschlagene Angebot eingehen. Ein Penny ist immerhin besser als gar nichts. Und beide Seiten wissen das. Aber so verhalten wir Menschen uns nicht. Im Mittel liegen die Angebote im Ultimatumspiel ein gutes Stück höher als im Diktatorspiel, nämlich zwischen 30 und 50 % (Camerer 2003). Angebote unter 20 % werden normalerweise ausgeschlagen. Das bedeutet: Den Empfängern ist es lieber, dass gar keiner einen Penny bekommt, als ein Angebot anzunehmen, das sie als zu niedrig erachten. Einmal mehr zeigt sich, dass das menschliche Verhalten in diesen Spielen von einer *Norm des Eigeninteresses* ebenso wie einer *Norm der Fairness* motiviert wird (Eckel & Grossmann 1995; Rabin 1993).

Doch verhalten wir uns immer nach dem Prinzip der Fairness, will heißen, machen wir immer halbe-halbe? Das kann man so nicht sagen. Die beiden Wirtschaftsforscher Van Dijk und Vermunt führten 2000 das Diktator- und das Ultimatumspiel unter den Bedingungen der symmetrischen Information (Spieler und Gegenspieler haben die gleichen Informationen) und der asymmetrischen Information (ein Spieler besitzt Informationen, die der andere nicht hat) durch. Die Manipulation der Informationen ist

äußerst relevant, und zwar insofern, als auf Seiten der Diktatoren wie auch der Vorschlagenden jede Geldeinheit den doppelten Wert hat, nur dass die jeweiligen Gegenspieler im Austausch lediglich den einfachen Wert erhalten. Beispiel: Für eine Geldeinheit im Wert von 1 Dollar erhielten Diktatoren und Vorschlagende bei Auszahlung 2 Dollar, ihre Gegenspieler 1 Dollar. Die Diktatoren zeigten sich von dieser Information unberührt und machten unter beiden Bedingungen in etwa die gleichen Angebote. Die Vorschlagenden dagegen wurden von dieser Information sehr stark beeinflusst; sie machten Angebote, die das Geld in etwa halbe-halbe verteilten, wenn ihre Gegenspieler den wahren Wert der Geldeinheiten kannten, nutzten deren Unkenntnis unter den Bedingungen der asymmetrischen Information aber maßlos aus, indem sie ihnen ein nur scheinbar faires Halbe-halbe-Angebot machten. In der Realität würde eine solch scheinbar gleiche Verteilung bedeuten, dass der Vorschlagende einseitig ein Drittel mehr Geld einstriche als sein Gegenspieler. Die Diktatoren hingegen hatten ebenfalls mehr Macht und mehr Information als ihre Gegenspieler, verhielten sich ihnen gegenüber aber sehr viel fairer. Es war, als habe überhaupt erst dieser große Informationsvorsprung den Anstoß zu einer Norm der Fairness gegeben. Die Vorschlagenden im Ultimatumspiel nutzen jedoch den Informationsvorsprung gegenüber ihren Gegenspielern aus und verhielten sich egoistisch. Oder, wie ein Wirtschaftswissenschaftler sagen würde, sie verhielten sich in der Vorteilssituation der asymmetrischen Information strategisch – wie Börsenmakler, die sich an Insidergeschäften beteiligen, oder Energieriesen, die ihre Mitarbeiter ermuntern, ihre Aktien zu halten, während sie ihre eigenen bereits verkau-

fen, wohlwissend, dass das Unternehmen angeschlagen ist und die Aktien ohnehin bald wertlos sind.

Ein ganz anderer methodischer Ansatz von Fiddick und Cummins (2007) lieferte noch weitere, sehr interessante Ergebnisse. Die Teilnehmer wurden gebeten, ein Fahrgemeinschaftsprojekt zu bewerten, bei welchem der eine das Benzingeld, der andere das Fahren übernimmt. Man legte ihnen hypothetische Bücher vor, die im Verhalten der Benzingeldzahler eine starke Variabilität zeigten (sie erfüllten ihren Teil der Abmachung zwischen 100 und gerade einmal 25 %). Anschließend wurden die Teilnehmer gefragt, ob sie das Projekt unter konsequenter Einhaltung der Bedingungen fortführen wollten, und wie fair sie das Verhalten ihres Gegenspielers einschätzten. Das Besondere an diesem Experiment war, dass es einige Szenarien gab, in denen die beiden Akteure den gleichen Status und die gleiche Macht hatten (beide waren Angestellte), und einige, in denen Status und Macht ungleich verteilt waren (einer der Akteure war Chef des anderen). Die Teilnehmer zeigten sich sehr viel toleranter, wenn der Angestellte den Chef beschummelte als umgekehrt der Chef den Angestellten. Dies war auch dann der Fall, wenn der Angestellte sehr viel mehr Geld einstrich als der Chef. Damit diese Asymmetrien *in puncto* Toleranz und empfundener Fairness entstehen konnten, musste ein wichtiger Faktor gegeben sein: Der Angestellte musste unmittelbar der Angestellte des Chefs sein. Wenn Chef und Angestellter unterschiedlichen Firmen in derselben Branche angehörten, war dieser Effekt nicht feststellbar; vielmehr zeigte sich ein gleiches Maß der Intoleranz für betrügerisches Verhalten unabhängig vom Beschäftigungsstatus. Es sieht also ganz danach aus, als sei-

en es die Asymmetrien der sozialen Beziehungen und nicht die Asymmetrien der Kosten und persönliche Vorteile, die diesem Effekt unterliegen.

Fiddick und Cummins bezeichneten diese Ergebnisse als den *noblesse oblige*-Effekt, der hier dem Sinn nach so viel bedeutet wie „Status und Macht verpflichten" oder „mit Reichtum, Macht und Erfolg erwächst Verantwortung gegenüber den vom Glück weniger Begünstigten". In der Ethik wird dieser Begriff bisweilen gebraucht, um eine moralische Wirtschaft zu beschreiben, in der Privilegien ausgewogen (symmetriert) werden durch die moralische Verpflichtung eines Individuums gegenüber all jenen, die keine Privilegien haben. Das großzügige Verhalten der Diktatoren in der Studie von Vermunt und van Dijk (2000) kann ebenfalls in dieser Weise beschrieben werden. Auch die Diktatoren verfügten alleinig über sowohl die Geldwerte als auch die wichtigste Information, nutzten die Situation aber dennoch nicht zu ihrem Vorteil aus. Sie verhielten sich vielmehr so, als würde sie dieses Privileg zu einer geradezu pastoralen Verantwortung gegenüber ihren Gegenspielern leiten.

Sehen wir uns nun an, was passiert, wenn die Spieler sich allein aufgrund der jeweiligen Konkurrenzfähigkeit einschätzen. In einer Reihe von Verhaltensexperimenten führten Hoffman et al. (1994) mit den Teilnehmern im Vorfeld ein kleines Konkurrenzspiel durch, um herauszufinden, ob der relative Status einer Person das Verhalten der Spieler im Diktator- und Ultimatumspiel relevant beeinflusst. Die Teilnehmer wurden gebeten, an einem kleinen Fragespiel zu aktuellen Ereignissen aus dem Tagesgeschehen teilzunehmen. Anschließend wurden sie namentlich nach der Zahl ihrer richtig gegebenen Antworten in einer

Bestenliste platziert. Diese Liste wurde für alle Spieler einsehbar ausgehängt. Danach stellten die Experimentatoren die Teilnehmer paarweise zusammen, und zwar immer eine höherrangige mit einer niederrangigeren Person. Die Person, die Rang 1 einnahm, kam mit der Person auf Rang 10 zusammen, die Person auf Rang 2 mit der auf Rang 11, die Person auf Rang 3 mit der auf Platz 12 und so weiter. Der jeweils höherrangigen Person in einem Paar wurde im Diktatorspiel die Rolle des Diktators zugewiesen, im Ultimatumspiel die Rolle des Vorschlagenden.

Die Ergebnisse waren erstaunlich. Im Vergleich zu einer Kontrollgruppe, die man im Vorfeld keinem Konkurrenzspiel unterzog, nahmen die Diktatoren sehr ungleiche Verteilungen vor (zu ihren Gunsten). Die Vorschlagenden hingegen machten deutlich niedrigere Angebote, ohne dass es dadurch zu einer Steigerung der Ablehnungsrate gekommen wäre (auch hier wieder verglichen mit einer Kontrollgruppe ohne ein Konkurrenzspiel im Vorfeld). Mit anderen Worten: Die höherrangigen Personen glaubten, sie hätten Anspruch auf mehr, während die niederrangigen glaubten, es stünde ihnen weniger zu.

Die Neurowissenschaft zeigt, warum wir uns verhalten, wie wir uns verhalten

Ziehen wir einmal Bilanz: Wie die Ergebnisse der experimentellen Studien zeigen, verhalten wir uns als Entscheider generell weniger egoistisch und weniger strategisch, und sie offenbaren eine größere Bedeutung sozialer Aspekte wie

Reziprozität, Fairness und den relativen sozialen Status einer Person, als es die Spieltheoretiker prognostizieren. Aber woran liegt das? Liegt es einfach nur daran, dass wir uns ständig irren? Oder versuchen wir uns in einer Weise zu verhalten, die ein Spieltheoretiker als rational definieren würde, hinter der wir aufgrund der menschlichen Fehlbarkeit aber immer zurückbleiben? Oder ticken wir einfach so? Ja, genau das. Wir ticken einfach so. Das jedenfalls legen die Ergebnisse neurowissenschaftlicher Studien nahe.

Wie wir im folgenden Kapitel noch ausführlicher besprechen werden, laufen Wahrscheinlichkeitsinformationen, die das Entscheidungsverhalten relevant mitbestimmen, in einem vorderen Hirnareal zusammen. Gleichwohl ist die Aktivierung tieferer Hirnareale abhängig vom Maß oder Wert eines Stimulus wie Belohnung oder Bestrafung. Belohnung oder Bestrafung können monetärer oder sozialer Natur sein. Was also passiert in unserem Gehirn? Um dies darzustellen, führte man Studien durch, in denen die Teilnehmer das Gefangenendilemma spielten und dabei gleichzeitig eine funktionelle Magnetresonanztomografie (fMRT oder fMRI) durchgeführt wurde, die Bilder der aktivierten Hirnareale lieferte. Kurzum: In der Interaktion mit einem Spielpartner führt die erwiderte Kooperation (positive Reziprozität) immer zu einer erhöhten Aktivierung im Striatum (Belohnungsareal), ganz egal, ob die gleiche Summe Geld gewonnen oder verloren wird; die nichterwiderte Kooperation (negative Reziprozität) hingegen zeigt einen entsprechenden Rückgang der Aktivierung in diesem Hirnareal (Sanfey 2007). Diese Ergebnisse lassen darauf schließen, dass wir in Gefangenenspielen Kooperation mit Belohnung erwidern, fehlende Kooperation hingegen mit Bestrafung.

Rilling et al. (2002) fassen diese Erkenntnisse wie folgt zusammen: Eine Aktivierung jener Schaltkreise, die Teil des Belohnungssystems im Gehirn sind, verstärkt die Kooperationsbereitschaft positiv und motiviert den Einzelnen, sich nicht zu einer egoistischen Defektion hinreißen zu lassen.

Einschlägige Studien hierzu ergaben, dass die neuronalen Belohnungssysteme aktiviert werden, wenn wir selbst Geld für wohltätige Zwecke spenden oder wir unmittelbar wahrnehmen, dass Geld für wohltätige Zwecke gespendet wird. Dies gilt aber nur dann, wenn die Spenden freiwillig erfolgen; sind sie unfreiwillig, weil es irgendein äußerer Umstand verlangt, leuchten die Belohnungszentren im Gehirn nicht auf. Das bedeutet, dass wir es als lohnenswert erachten, uns altruistisch (und nicht nur kooperativ) zu verhalten.

Wir wissen, dass dritte Personen, die ein Gefangenendilemma lediglich beobachten und gar nicht selbst mitspielen, wild entschlossen sind, defektierende Spieler zu bestrafen, selbst ihnen persönlich dadurch Kosten entstehen. Wie sich herausstellt, kommt es zu einer Aktivierung der Belohnungszentren im Gehirn eines Spielers, wenn er die Möglichkeit bekommt, den Defektor zu bestrafen, selbst wenn er dabei Geld verliert. Dies bedeutet, dass wir es als lohnenswert erachten, defektierende Personen zu bestrafen. Unter der Voraussetzung, dass Öffentliches-Gut-Spiele dieselbe mathematische und soziale Struktur haben wie Gefangenendilemma-Spiele, dürfte es nicht weiter verwundern, dass diese Spiele im Ergebnis die gleichen neuronalen Aktivitäten zeigen: Belohnungsabhängige Hirnareale werden aktiviert, wenn Trittbrettfahrer bestraft werden (de Quervain et al. 2004).

Im Vertrauensspiel war die Aktivität in den neuronalen Belohnungszentren am stärksten, wenn der Investor Groß-

zügigkeit mit Großzügigkeit erwiderte, und am schwächsten, wenn der Investor Großzügigkeit mit Geiz vergalt. Sehr viel verblüffender ist, dass der Geldbetrag, den der Investor zu zahlen bereit ist, chemisch manipuliert werden kann, und zwar mittels eines Hormons namens Oxytocin. Oxytocin wird gelegentlich auch als „Bindungshormon" bezeichnet, denn es scheint soziale Bindungen und Vertrauen zu fördern. Bei Bedarf wird es aus dem Hinterlappen (Neurohypophyse) der Hirnanhangdrüse (Hypophyse) abgegeben. Bei schwangeren Frauen bewirkt es wehenauslösende Kontraktionen der Gebärmutter und das Einschießen der Muttermilch in die Brust. Dies bedeutet, dass das Gehirn von Baby und Mutter während der Geburt mit Oxytocin geradezu überflutet ist und in ein Wohlgefühl taucht, das die (emotionale) Bindung zwischen beiden stärkt. Auch beim Sex setzt das Gehirn große Mengen an Oxytocin frei und verstärkt die emotionale Bindung, diesmal zwischen Mann und Frau. In einer Studie von Kosfeld et al. (2005) wurde den Investoren und Treuhändern vor Beginn des Vertrauensspiels entweder Oxytocin oder ein Placebo verabreicht. Das Oxytocin steigerte die Vertrauensbereitschaft der Investoren – sie händigten dem Treuhänder mehr Geld aus. Auf das Verhalten der Treuhänder wirkte sich dies allerdings nicht aus. Sie verhielten sich so, wie sich Treuhänder in diesen Studien immer verhalten, zahlten also etwas weniger zurück als der Investor übergeben hatte. Das Oxytocin steigert also die Vertrauensbereitschaft, nicht aber die Großzügigkeit. Ähnliche Ergebnisse zeigten sich im Ultimatumspiel: Intranasal verabreichte Oxytocingaben steigerten die Großzügigkeit um 80 %, hatten aber keine Auswirkungen auf das Verhalten der Diktatoren in diesem Spiel (Zak et al.

2007). Einmal mehr zeigt sich, dass uns das Oxytocin zwar vertrauensvoller macht, aber nicht zwangsläufig auch großzügiger. Und was empfinden wir bei geizigen bzw. großzügigen Angeboten? Im Ultimatumspiel aktivieren geizige Angebote Gehirnareale (die Insula anterior), die mit Gefühlen der Abscheu verbunden sind (Sanfey et al. 2003). Doch dies gilt nur, wenn ein Spieler einen menschlichen Gegenspieler hat; wenn der Gegenspieler ein Computer ist, treten in den entsprechenden Gehirnarealen keine Veränderungen auf.

Die Ergebnisse neurowissenschaftlicher Entscheidungsstudien wie diesen zeigen, dass der Einfluss sozialer Aspekte auf diese Spiele nicht hoch genug eingeschätzt werden kann. Wir verhalten uns, als seien wir auf reziproke Langzeitbeziehungen gepolt. Belohnungszentren werden aktiviert, wenn wir kooperativ und großzügig sind, und die Hirnareale, die negative Reize wie Abscheu steuern, werden aktiviert, wenn wir egoistischen Verhaltensweisen begegnen. All diese neuronalen Schaltkreise scheinen darauf programmiert, ein bestimmtes Sozialverhalten zu erzeugen: Ungleichheiten möglichst zu vermeiden suchen, das Prinzip der Reziprozität zu fördern sowie auf Bestrafungen all jener zu drängen, die sich auf Kosten anderer einen Vorteil verschaffen. In wiederholten Spielen – in wiederholten Transaktionen zwischen Individuen, die sich an ihren jeweiligen Transaktionspartner und an die Transaktionshistorie erinnern werden – wird die Reputation zu einem überaus wichtigen Faktor. Wer will schon gerne Transaktionen vornehmen mit Personen, die im Ruf stehen, geizig und egoistisch zu sein – schon gar nicht mit welchen, die möglicherweise planen, sich in solch einer Weise zu verhalten! Ganz gleich, ob man

selbst egoistisch oder großzügig ist, Transaktionen mit einer kooperativen Person sind immer lohnender.

Tatsächlich zeigt sich genau dieser Punkt auch bei Kleinkindern. In einer Reihe von Studien (Hamlin et al. 2007) beobachteten Kinder, die erst sechs Monate alt waren, einen animierten roten Kreis, der sich mühsam einen steilen Hang hinaufbewegt. In der ersten Situation taucht ein gelbes Dreieck auf und schiebt den roten Kreis bis oben auf den Gipfel hinauf. In der zweiten Situation taucht ein blaues Viereck auf und schiebt den roten Kreis den Hang wieder hinunter. In weiteren Situationen taucht ein drittes Objekt auf, das dem Kreis entweder nicht behilflich ist oder das nicht animiert war und sich folglich nicht bewegte. Nachdem die Kinder den kleinen Film gesehen hatten, zeigte man ihnen die drei Objekte und ließ sie aussuchen, mit welchem sie spielen mochten. Die Kinder zogen in der überwiegenden Mehrzahl die Helfer (Kooperanden) den anderen Objekten, die nicht halfen oder sich nicht bewegt hatten, vor. Die Autoren der Studie folgerten erstens daraus, dass selbst vorsprachliche Kleinkinder einen individuellen Akteur aufgrund seines Verhaltens gegenüber einem anderen bewerten, und zweitens, dass es sich bei dieser Art der sozialen Evaluation um eine biologische Adaptation handelt.

Die Evolution der Kooperation

Neurowissenschaftliche Studien zum Entscheidungsverhalten scheinen zu belegen, dass das menschliche Verhalten auf Kooperation ausgelegt ist. Und evolutionsbiologische Studien scheinen zu belegen, dass es nicht nur darauf aus-

gelegt, sondern in unseren Genen angelegt, sprich genetisch codiert ist. Kooperation ist uns quasi angeboren. Doch mit dieser einfachen Erklärung gibt sich die Forscherwelt nicht zufrieden. Sie will wissen, *warum* uns die Kooperation angeboren ist. Die aufschlussreichste Antwort darauf kommt aus der Evolutionsbiologie.

Axelrod und Hamilton stellen in ihrem klassischen Aufsatz über das 1981 veranstaltete und unter Computerprogrammen ausgetragene Turnier zum Gefangenendilemma heraus: „Die Evolutionstheorie basiert auf dem Kampf um Leben und das Überleben des Stärkeren. Die Kooperation jedoch ist unter Mitgliedern derselben Art und selbst unter Mitgliedern verschiedener Arten allgemein üblich." Das Phänomen der Kooperation ist mit der Theorie vom egoistischen Gen als evolutionsbiologische Grundlage nur schwer vereinbar – ein Theorie, die besagt, dass der Prozess der Evolution letztlich nur die genetische Information, die dem eigenen Interesse und damit dem Erhalt der eigenen Art dient, weiterzugeben sucht (Dawkins 1976). Wenn wir egoistisch handeln (beispielsweise unsere Lebensmittel nur für uns behalten), erhöhen wir unsere Überlebenschancen. Und wenn wir unsere Überlebenschancen erhöhen, ermöglichen wir uns ein längeres Leben. Und wenn wir länger leben, haben wir mehr Möglichkeiten uns fortzupflanzen. Folglich bleiben unsere Gene in der menschlichen Population länger erhalten. Damit wird klar, wie Gene, die egoistisches Verhalten fördern, sich erfolgreich durchsetzen können. Aber wie setzen sich Gene durch, die das nichtegoistische Verhalten und damit Formen der Kooperation begünstigen?

Hamilton gibt darauf folgende Antwort: Verwandtenselektion. Die Verwandtenselektion lässt sich in einfachen Worten wie folgt erklären: Die Gene, die ein Individuum mit einem anderen teilt, setzen sich nur dann erfolgreich durch, wenn der Nutzen der Genweitergabe größer ist als die Kosten, die dafür aufgebracht werden. Mal angenommen, sie haben eine Menge Geld und ihre Schwester hat alle Mühe, sich und ihren kleinen Sohn durchzubringen. Wenn Sie ihr nun etwas von Ihrem Geld abgeben, um ihr das Leben mit Ihrem Neffen zu erleichtern, haben Sie das Überleben Ihrer Blutsverwandten und damit das Fortleben der gemeinsamen Gene in künftigen Generationen gesichert, ohne dass es sie groß etwas gekostet hätte. Insofern folgt das Motiv, den eigenen Verwandten zu helfen, durchaus der Theorie vom egoistischen Gen.

Aber wie erklärt sich die Kooperation zwischen nichtverwandten Individuen? In der Natur kommt es sehr häufig zu kooperativen Verhaltensweisen. Nehmen wir als Beispiel die Putzerfische: Sie schwimmen ins Maulinnere von anderen Fischen, weiden dort Parasiten und abgestorbene Hautteilchen ab und schwimmen durch Maul und Kiemendeckel wieder hinaus. Putzerfische können ihre „Kunden" gefahrlos anschwimmen und werden nicht gefressen. Die „Kunden" profitieren von der Körperpflege und der Putzerfisch profitiert von einer reichen Nahrungsquelle – ein Gewinn für beide Seiten. Dieses Prinzip findet sich in einer etwas komplexeren Ausgestaltung auch bei den Vampirfledermäusen. Sie sind im Grunde genommen Blutsauger, da sie sich vom Blut anderer Tiere ernähren. Diese Blutmahlzeit können sie mit anderen Fledermäusen teilen, tun dies für gewöhnlich aber nur mit jenen, die ihre Mahlzeit in

der Vergangenheit selbst wiederum mit ihnen geteilt haben (Wilkinson 1984), und zwar unabhängig davon, ob sie mit ihnen verwandt sind oder nicht. In der Welt der Primaten kommt das Prinzip der Reziprozität unter Nichtverwandten häufig vor. Grüne Meerkatzen beispielsweise reagieren auf Hilferufe anderer nichtverwandter Gruppenmitglieder besonders stark, wenn kurz zuvor eine gegenseitige Fellpflege stattgefunden hatte. Mit ihnen bilden sie auch die stärksten Allianzen (Cheney & Seyfarth 1992). Und genau wie menschliche Individuen in einem Öffentliches-Gut-Spiel rächen sich Schimpansen an Individuen, die ihre Nahrung nicht teilen, reagieren ihrerseits aggressiv (de Waal 1989), wenn diese Nahrung fordern, oder sie geben falsche oder gar keine Information über den Ort ihrer Nahrungsquelle (Woodruff & Premack 1979).

Wie diese Beispiele zeigen, kommt die Kooperation beiden Seiten zugute und begünstigt folglich die Genweitergabe. Warum aber sollte die natürliche Selektion das altruistische Verhalten eines anderen, nichtverwandten Individuums begünstigen und nicht das eigene, zumal sich dieses Individuum bei jeder Transaktion durch Defektion einen noch größeren Vorteil verschaffen könnte? Warum frisst der „Kunde" den Putzerfisch nicht einfach auf, nachdem der seine Arbeit getan hat? Warum teilen Vampirfledermäuse und Primaten ihre Mahlzeiten und nehmen Kosten auf sich, um jene zu bestrafen, die nicht teilen?

Eine maßgebende Lösung für all diese Rätsel offerierte 1971 der US-amerikanische Evolutionsbiologe Robert Trivers mit seinem Konzept des reziproken Altruismus (Box 2.2). Er zeigte, dass altruistisches Verhalten selektiv entstehen kann, wenn beide Seiten, Hilfeleister wie Hilfe-

empfänger, einen Nutzen davon haben. Er schreibt: „Jeder hilft dem anderen, während jeder sich gleichzeitig selbst hilft." Das Problem des reziproken Altruismus besteht darin, dass ein kooperierendes Verhalten für den Einzelnen von Nutzen sein kann, er die Erwartung der Reziprozität aber auch ausnutzen kann. Trivers zeigte, dass die Selektion Betrüger ausschließt, die durch empfangene Hilfe einen hohen Nutzen haben, diese Hilfe später aber nicht erwidern und ihr eigener Nutzen somit überwiegt. Wann kommt es dazu? Es kommt dann dazu, wenn ein Helfer auf einen Betrüger reagiert, indem er ihn von künftigen Transaktionen ausschließt. Unter diesen Bedingungen ist der reziproke Altruismus eine evolutionär stabile Strategie (ESS). Das Konzept der ESS wurde von John Maynard Smith 1972 eingeführt. Eine Strategie wird als evolutionär stabil bezeichnet, wenn eine Population von Individuen sie anwendet und damit eine kleine Subpopulation, die eine andere Strategie anwendet, übertreffen und eliminieren kann.

Box 2.2 Die Mathematik des reziproken Altruismus – oder: Das egoistische Gen

Betrachten wir zwei Genpopulationen, die Altruisten (A) und die Nicht-Altruisten (N). A verhält sich altruistisch, wenn der Gesamtnutzen des reziprok altruistischen Verhaltens seine Gesamtkosten übersteigt.

Kostenreduktion bei Weitergabe der Gene

Nutzenzuwachs bei Weitergabe eigener Gene

Angenommen, das altruistische Verhalten wird gesteuert durch das Allel (a_2) eines Gens an einer bestimmten Stelle im Genom (Locus). Neben diesem gibt es ein weiteres Allel (a_1) dieses Gens, das sich an der gleichen Stelle im Genom befindet, sich aber darin von a_2 unterscheidet, dass

es ein nichtaltruistisches Verhalten befördert. Wie wird sich das altruistische Verhalten verbreiten?

Zufallsverteilung des Altruismus

Drei Genotypen sind möglich: a_1a_1, a_1a_2, a_2a_2

a_1a_1 (Nicht-Altruist): profitiert von $(1/N)\Sigma b_i$, wobei b für den Empfänger von Nutzen ist.

a_2a_2 (Altruist): hat einen Nettonutzen von $(1/N)\Sigma b_i - (1/N)\Sigma c_j$, wobei c die Kosten für a_2a_2 (Geber) darstellen.

Ergebnis: $(1/N)\Sigma c_j < 0$, also wird a_1 das Allel a_2 in der Population ersetzen.

Nicht-Zufallsverteilung des Altruismus (nur bezogen auf Verwandte)

Der Altruismus wird sich verbreiten, wenn der Gesamtnutzen für den Empfänger den Gesamtnutzen für den Geber erheblich übersteigt (a_2 wird a_1 ersetzen) (Hamilton 1964) $r > c/b$

Nicht-Zufallsverteilung des Altruismus (nur bezogen auf Reziprokatoren)

Der Altruismus wird selektiert, wenn der Nettonutzen für den Altruisten a_2a_2 den Nutzen für den Nicht-Altruisten a_1a_1 übersteigt.

$(1/p^2)(\Sigma b_k - \Sigma c_j) > (1/q^2)\Sigma b_m$ wobei

b_k = Nutzen für den Empfänger a_2a_2

c_j = Kosten für den Geber a_2a_2

b_m = Nutzen für den Empfänger a_1a_1

p = Häufigkeit des Allels a_2

q = Häufigkeit des Allels a_1

Notwendige Bedingung: Σb_m muss gering bleiben

Das passiert, wenn der Altruist ein betrügerisches Verhalten seitens des Empfängers erwidert, indem er künftige Transaktionen an eben jenen kürzt.

Wird der reziproke Altruismus wie im Gefangenendilemma modelliert, dann hat die Auszahlungsmatrix folgende Beschränkungen: $T > R > P > S$ und $R > (S + T)/_2$. Dies bedeutet, dass die Versuchung zu defektieren größer sein muss, als die Belohnung für ein kooperatives Verhalten, was wiederum größer sein muss als die Bestrafung für ein nichtko-

operatives Verhalten. Die Auszahlung für den Defektor fällt am geringsten aus – mit einem Defektor zu kooperieren, geht also nach hinten los. Am Ende sollte die Belohnung größer sein als die Hälfte der Auszahlungssumme an den Defektor und größer als bei einer Versuchung zu defektieren. Diese Form der Matrix bildet das Szenario, das dem Gefangenendilemma zugrunde, liegt am besten ab.

Etliche Forscher führten Studien dazu durch. Sie modellierten das Prinzip des reziproken Altruismus als ein Gefangenendilemma, und zwar so, dass es den von Trivers identifizierten Bedingungen entsprach. In einem einmaligen Gefangenendilemma ist die Defektion eine ESS. In wiederholten Spielen aber ist die Kooperation eine ESS, und zwar dann, wenn die Teilnehmer sich gegenseitig als Betrüger erkennen und von künftigen kooperativen Handlungen ausschließen können.

Im Klartext bedeutet das: Wenn Sie die andere Person nie zuvor gesehen haben, nichts über sie wissen und auch keine gemeinsame Geschichte mit ihr haben, und wenn keiner von Ihnen den anderen jemals wiedersehen wird, dann ist die Strategie mit dem größten Nutzen die, sich das Geld zu schnappen und aus dem Staub zu machen. Wenn Sie die andere Person aber kennen, ihre Reputation kennen oder eine gemeinsame Geschichte haben, und Sie beide wissen, dass Sie sich auch künftig noch öfter begegnen werden, wird sich letztlich die Kooperation durchsetzen. Allerdings nur dann, wenn Sie sich fernhalten von Betrügern – von jenen also, die sich das Geld schnappen und aus dem Staub machen. Denn wie die Simulationen durchweg zeigen, kann der wechselseitige Altruismus nicht dauerhaft

aufrechterhalten werden und erlischt, wenn Sie fortlaufend Transaktionen mit Betrügern tätigen (Sober & Sloan-Wilson 1999). Insofern müssen Betrüger von künftigen Transaktionen ausgeschlossen werden; andernfalls verschwindet die Kooperation aus der Population.

Trivers führt außerdem folgende Merkmale auf, die die Evolution der Kooperation begünstigen: lange Lebenszeiten der Individuen in einer Population, niedrige Dispersionsrate (die Menschen bleiben am Ort), wechselseitige Beziehungen unter den Mitgliedern der Population, hohes Maß an elterlicher Fürsorge (welche Verwandtenselektion und Altruismus begünstigt), Hilfeleistungen in Verteidigungs- und Kampfsituationen, Fehlen einer streng linearen Dominanzhierarchie. Wie sich herausstellt, trifft diese Beschreibung haargenau auf die Jäger-und-Sammler-Kulturen in der Frühzeit der Menschen zu. Nach Trivers hat im Laufe der menschlichen Evolution eine scharfe Selektion zugunsten reziprok altruistischer Verhaltensweisen, sprich der Kooperation stattgefunden. Individuen früher Populationen, die besser kooperierten, setzten sich langfristig durch. Sie lebten länger und pflanzten sich stärker fort als jene, die sich nicht kooperativ verhielten und gaben ihre altruistischen Gene damit über viele Generationen weiter. Und genau das ist der Grund, *warum* wir so geboren werden. Genau darum tätigen wir soziale oder monetäre Transaktionen mit einer ausgeprägten Neigung zur Kooperation. Genau darum erwarten wir Reziprozität und sind bereit zu zahlen, um Egoisten für nichtkooperatives Verhalten zu bestrafen. Wir sind genetisch codiert, die Kooperation aufrechtzuerhalten, denn nur wenn wir kooperieren, können wir langfristig überleben.

3

Rationale Entscheidung

Wir wählen die Handlungsalternative aus, die unsere gewünschten Ziele weitestmöglich realisiert

In seinem 2002 erschienen Buch *Calculated risks* berichtet Gerd Gigerenzer von einem Arzt, der 90 Frauen mit angeblich hohem Brustkrebsrisiko überzeugen konnte, sich ihre gesunden Brüste vorsorglich abnehmen zu lassen im guten Glauben „ihre Brüste zu opfern in einem heroischen Tausch gegen die Gewissheit, ihr Leben zu retten und ihre Lieben vor Leid und Verlust zu bewahren". Doch hätte der Arzt die Zahlenstatistik richtig gelesen, so Gigerenzer, hätte er gewusst, dass für die große Mehrheit dieser Frauen gar nicht zu erwarten stand, an Brustkrebs zu erkranken. Was hat es damit auf sich?

Die American Medical Association (AMM), die größte Standesvertretung der Ärzte in den USA, änderte unlängst die Bestimmungen für das Brustkrebs- und Prostata-Screening mit der Begründung, dass häufige Untersuchungen zu viele falsch-positive Testergebnisse erbrächten (Ergebnisse, die eine Krebserkrankung befinden, wenn in Wirklichkeit keine vorhanden ist). Die AMM gab die Empfehlung, sich seltener untersuchen zu lassen. Ein Aufschrei ging durch die Reihen der Patienten. TV-Experten behaupteten, hinter der

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_3,

Entscheidung stecke allein die Absicht, die medizinischen Kosten zu reduzieren, und das wiederum würde bedeuten, dass die Ärzteschaft bereit sei, Leben zu opfern, um Geld zu sparen. Hatte diese Entscheidung also medizinische Gründe oder finanzielle, oder ist sie nur ein weiteres Beispiel für Fehlurteile im täglichen Denken? Dieses Kapitel wird diese Frage erhellen.

Wie „große Entscheidungsmacher" ihre Entscheidungen treffen

Eckstein der klassischen Entscheidungstheorie ist das Konzept eines rationalen Akteurs. Doch was der Durchschnittsbürger unter „rational" versteht, ist nicht unbedingt dasselbe wie das, was ein Wirtschaftswissenschaftler damit meint. Man mag einen Menschen als rational beschreiben, wenn seine Denkweise besonnen, logisch und schlüssig ist. Für einen Wirtschaftswissenschaftler ist ein rationaler Akteur in erster Linie eigennützig – das heißt, er vergleicht vorhandene Optionen und wählt diejenige aus, die den eigenen Nutzen maximiert und die eigenen Kosten minimiert.

Die *Theorie der rationalen Entscheidung* kann wie folgt zusammengefasst werden: Wir treffen rationale Entscheidungen, indem wir mögliche Ereignisse nach der Wahrscheinlichkeit ihres Eintretens bestimmen und sie mit ihrem jeweiligen Wert multiplizieren (Edwards 1955). Das optimalste Ergebnis ist dasjenige mit dem größten Produkt – das heißt, wir präferieren die Option, die für uns mit dem maximalen Wert einhergeht, also die, mit der wir unsere gewünschten Ziele weitestmöglich realisieren können.

In einem rationalen Entscheidungsprozess sind danach zwei Faktoren zu berücksichtigen: Der erste bestimmt die Wahrscheinlichkeit des Eintretens eines erwünschten Ergebnisses. Der zweite bestimmt den Wert, den dieses Ergebnis für sie hat. Einfaches Beispiel: Mit wem würden Sie sich lieber verabreden? Mit Brad Pitt, mit dem netten, durchschnittlich aussehenden Jungen von nebenan oder mit dem langweiligen Typen, mit dem Sie Ihre Mutter ständig verkuppeln will? Brad Pitt ist ganz klar der bestaussehende von allen und führt ein ziemlich aufregendes Leben. Sich mit ihm zu verabreden, ist also erstrebenswerter, als sich mit dem durchschnittlichen Jungen von nebenan zu treffen, und ein Date mit dem langweiligen Typen kommt am wenigsten infrage. Nun ist ein Date mit Brad nicht sehr wahrscheinlich, ein Date mit dem Nachbarsjungen schon eher und ein Date mit dem Langweiler kommt Ihnen erst gar nicht in die Tüte. Unterm Strich ist ein Date mit dem netten Jungen von nebenan also die beste Option: Eine Verabredung mit ihm wird eher eintreten als ein Date mit Brad Pitt (das Sie sich zweifelsohne eher wünschen würden) und ist allemal lohnender als ein Date mit dem von Ihrer Mutter favorisierten Langweiler (ganz klar!). Und siehe da, schon haben Sie eine rationale Entscheidung getroffen.

Der erwartete/erwünschte Wert eines möglichen Ergebnisses wird als der *Nutzen* einer Option bezeichnet. Er ist ein Maß für die Befriedigung (oder Freude). 1000 Dollar zu gewinnen, ist sichtlich befriedigender, als 100 Dollar zu gewinnen, weshalb wir lieber 1000 Dollar gewinnen als 100. Die Theorie der rationalen Entscheidung stellt zwei Annahmen für die Handlungspräferenz eines rationalen Individuums auf. Die erste besagt, dass eine „Vollständigkeit“

(*completeness*) gegeben sein muss, das heißt, der rationale Entscheider kann alle möglichen Handlungen in einer Reihenfolge nach ihrer Präferenz einstufen (auch dann, wenn er zwischen zwei oder mehreren Handlungsoptionen indifferent bzw. unentschieden ist). Die zweite Annahme über Präferenzen ist die „Transitivität" (*transitivity*), das heißt, wenn der Entscheider drei Optionen bewertet und Option 1 gegenüber Option 2 bevorzugt, und Option 2 gegenüber Option 3, dann zieht er auch Option 1 gegenüber Option 3 vor. Diesen beiden Annahmen zufolge kann ein rationaler Entscheider seine Optionen also nach Präferenz ordnen und schließen, dass seine Präferenzen in sich konsistent sind. Um nun auf unser Beispiel mit dem potenziellen Date zurückzukommen, könnten wir die Präferenzen wie folgt ordnen: Brad > Junge von nebenan > Langweiler. Würde ich Sie also vor die Wahl zwischen Brad Pitt und dem Jungen von nebenan stellen, müssten Sie sich demnach für Brad entscheiden. Wenn ich Sie vor die Wahl zwischen dem Jungen von nebenan und dem Langweiler stellte, müssten Sie sich demnach für den Jungen von nebenan entscheiden. Und stellte ich Sie schließlich vor die Wahl zwischen Brad Pitt und dem Langweiler, fiele Ihre Entscheidung ebenfalls auf Brad. Wenn nicht, verletzten Ihre Entscheidungen die Norm der Transitivität und ich hätte keinerlei Ahnung, wie ich Ihre Entscheidungen vorhersagen könnte.

Das Konzept vom Nutzen ist für gewöhnlich jedem sofort klar: Es ist das, was einen glücklicher und zufriedener macht. Etwas schwieriger zu verstehen, ist der zweite Aspekt, der die Theorie der rationalen Entscheidung ausmacht – die Berechnung der Wahrscheinlichkeit. In unserem Beispiel mit den drei potenziellen Dates haben wir

geschätzt, wie wahrscheinlich es ist, ein Date mit einem dieser drei Männer zu bekommen. Aber was, wenn man Sie beispielsweise auffordern würde, die Wahrscheinlichkeit einzuschätzen, dass Sie tatsächlich an Krebs erkrankt sind, sollten Sie nach einer Mammografie oder einem Test auf das prostataspezifische Antigen (PSA) einen positiven Befund für Brust- oder Prostatakrebs erhalten? Um diese Frage zu beantworten, brauchen wir einige Informationen, die in den Berichten zur Krebsstatistik (Surveillance, Epidemiology, and End Results, SEER), erhoben vom US-amerikanischen Krebsforschungszentrum NCI (National Cancer Institute), zu finden sind.

Beginnen wir mit dem Brustkrebs. Aus diesen Berichten geht hervor, dass 12 % der heute geborenen Frauen im Laufe ihres Lebens irgendwann einmal die Diagnose Brustkrebs bekommen werden. Man spricht hier vom Lebenszeitrisiko, Krebs zu entwickeln, das auf der sogenannten Krebsinzidenzrate basiert, das heißt auf der Zahl an Brustkrebsneuerkrankungen, die in den USA innerhalb eines Jahres aufgetreten sind. Diese Zahl beziffert die Neuerkrankungen pro 100.000 Menschen. Die Inzidenzrate ist zu unterscheiden von der Prävalenzrate, der Anzahl an Menschen, die in einer bestimmten Population zu einem bestimmten Untersuchungszeitpunkt an einer bestimmten Krebsart erkrankt sind. Die Inzidenzrate beziffert also die geschätzte Anzahl der Neuerkrankungen und die Prävalenzrate die Anzahl aller erfassten Fälle zu einem bestimmten Zeitpunkt.

Mammografien stellen aber keine zuverlässigen Indikatoren für künftige Krebserkrankungen dar. Es kommen digitale Röntgengeräte zum Einsatz, die Aufnahmen der weiblichen Brust liefern, über die ein erfahrener Radiologe dann

Tab. 3.1 Statistiken für Brustkrebs und Mammografien nach dem Alter der Patientinnen.

Brustkrebs			Mammografie			
Alter	**Inzidenz (%)**	**Prävalenz (%)**	**Alter**	**positiv (%)**	**Sensitivität (%)**	**negativ (%)**
30–39	0,34	0,17	30–39	nicht verfügbar	nicht verfügbar	nicht verfügbar
40–49	1,45	0,89	40–49	18,2	88,2	81,9
50–59	2,38	2,18	50–59	22,1	90,9	78,3
60–69	3,45	3,75	60–69	19,3	88,8	81,5

befinden und entscheiden muss, ob ein Krebs erkennbar ist. Die *Sensitivität* der Testmethode entspricht dem Anteil der positiven Mammografien von allen Mammografien bei Frauen, die tatsächlich Brustkrebs haben (Richtig-positiv-Rate). Die *Spezifität* entspricht dem Anteil der negativen Mammografien von allen Mammografien bei Frauen, die tatsächlich keinen Brustkrebs haben (Richtig-negativ-Rate). Doch Radiologen können Fehldiagnosen stellen. Ein falsch-positives Ergebnis bedeutet, dass der Radiologe entscheidet, eine Frau sei an Brustkrebs erkrankt, obwohl sie es tatsächlich gar nicht ist. Ein falsch-negatives Ergebnis bedeutet, dass der Radiologe entscheidet, eine Frau sei nicht an Brustkrebs erkrankt, obwohl sie es tatsächlich ist.

All diese Werte variieren als Funktion in Abhängigkeit vom Lebensalter. Tabelle 3.1 zeigt die Brustkrebsraten US-amerikanischer Frauen nach den Berichten zur Krebsstatistik (Surveillance, Epidemiology, and End Results, SEER) sowie nach den Angaben über die Mammografien aus der

US-Datenbank für Brustkrebs (Breast Cancer Surveillance Consortium; www.breastscreening.cancer.gov).

Man beachte, dass sowohl die Inzidenz- als auch die Prävalenzrate mit zunehmendem Alter ansteigen. Dies bedeutet, dass das Risiko, Brustkrebs zu entwickeln, mit dem Alter steigt. Auch andere Faktoren können das Erkrankungsrisiko erhöhen wie die Häufung von Brustkrebsfällen in der Familie oder Umweltfaktoren. Doch die angegebenen Zahlen zeigen, dass das Alter ein wesentlicher Faktor ist. Man beachte zudem, dass der Anteil der positiven Mammografien ebenfalls ansteigt, wohingegen der Anteil der richtig-positiv diagnostizierten Fälle (Sensitivität) und der Anteil der richtig-negativ diagnostizierten Fälle (Spezifität) sich nicht so stark verändert.

Angenommen, Sie unterzögen sich Ihrer ersten Mammografie und erhielten eine positive Brustkrebsdiagnose. Wie hoch ist die Wahrscheinlichkeit, dass Sie tatsächlich an Brustkrebs erkrankt sind? Nach den Zahlen der obigen Tabelle liegt die Wahrscheinlichkeit, dass eine 40-jährige Frau mit einer positiven Mammografie tatsächlich an Brustkrebs erkrankt ist, bei 5 %, für eine 50-jährige Frau bei 9 %.

Sind Sie überrascht von diesen Zahlen? Dachten Sie, sie müssten mindestens zehnmal höher liegen? Wenn ja, sind Sie damit nicht allein. Ihr Arzt könnte denselben Fehler gemacht haben: Im Rahmen einer Studie hatte man Ärzte gebeten, anhand ähnlicher Datenmengen die Wahrscheinlichkeiten von Brustkrebs zu ermitteln. Die Ergebnisse fielen düster aus (Eddy 1982). Die Schätzungen der Ärzte lagen um den Faktor 10 daneben – das heißt, wenn die tatsächliche Wahrscheinlichkeit, bei einem positivem Testergebnis tatsächlich an Brustkrebs erkrankt zu sein, bei 9 %

lag, so lagen die mittleren Schätzungen der befragten Ärzte bei 90 %! Es kann für Sie als informierter Patient daher nicht schaden, das Einmaleins dieser Zahlen zu kennen, um die richtigen Schlüsse zu ziehen.

Der sogenannte Satz von Bayes hilft Ihnen dabei. Thomas Bayes (1702–1761) war ein englischer Geistlicher, Laienmathematiker und leidenschaftlicher Spieler. Und als solcher war er natürlich brennend daran interessiert, wie sich die Chancen auf ein Gewinnerblatt im Kartenspiel, einen Gewinnerwurf im Würfelspiel oder das Gewinnerpferd beim Pferderennen ermitteln lassen (Box 3.1). In seinem Aufsatz *Essay towards solving a problem in the doctrine of chances* (der erst nach seinem Tod in Teilen von Richard Price überarbeitet und 1763 veröffentlicht wurde) schlägt er ein Entscheidungsverfahren vor, das auf Wahrscheinlichkeitsberechnungen basiert.

Box 3.1 Bayes und die Spielkarten

Lassen Sie uns nun all diese Ausführungen über Wahrscheinlichkeiten in einen Bereich übertragen, der den meisten von uns vertraut ist. Angenommen, ich ziehe eine Spielkarte aus einem normalen Kartendeck. Wie hoch ist die Wahrscheinlichkeit, dass ich einen König ziehe? Ganz einfach: Es gibt nur vier Könige im ganzen Deck und es gibt 52 Spielkarten. Also liegt die Wahrscheinlichkeit, einen König zu ziehen, bei 4:52. Dies ist die A-priori-Wahrscheinlichkeit, einen König zu ziehen, die wir hier als P(König) bezeichnen wollen.

Angenommen, Sie zögen eine Karte und ich sagte Ihnen, dass Sie tatsächlich einen König gezogen haben, stellte Ihnen aber die folgende Frage: Wie hoch ist die Wahrscheinlichkeit, dass Sie einen Karo-König gezogen haben? Nun, es gibt vier Könige und nur einer davon ist ein Karo-König. Also liegt die Wahrscheinlichkeit, dass einen Karo-König zu

ziehen, bei 1:4. Dies ist die A-posteriori-Wahrscheinlichkeit – die bedingte Wahrscheinlichkeit, dass die Karte ein Karo-König ist, vorausgesetzt ich weiß, dass die Karte ein König ist. Wir wollen sie hier als P(Karo|König) bezeichnen.

Lassen Sie uns nun das Spiel beginnen – will heißen, wir kombinieren Nützlichkeiten und Wahrscheinlichkeiten. Und wir spielen um Geld. Wenn Sie gewinnen, gewinnen Sie 10 Dollar. Was ist wahrscheinlicher, wenn ich eine Karte aus dem Deck ziehe?

A: Die Karte ist ein König.

B: Die Karte ist der Karo-König.

Kinderleicht, wenn man in Gewinnchancen oder Wahrscheinlichkeiten denkt: Es gibt vier Könige im Deck. Und es gibt nur einen Karo-König. Also verteilen sich die Gewinnchancen folgendermaßen:

$$\mathrm{A}: \mathrm{P}(\text{König}) = 4:52$$

$$\mathrm{B}: \mathrm{P}(\text{Karo} - \text{König}) = 1:52$$

A hat also die größte Wahrscheinlichkeit. Dieses Beispiel zeigt eine sehr wichtige Regel der Wahrscheinlichkeitstheorie – die Konjunktionsregel. Sie besagt, dass ein „Ereignis A niemals weniger wahrscheinlich sein kann als die Konjunktion der (unabhängigen) Ereignisse A und B".[1] Formal: $P(A\&B) \leq P(A)$. Angewendet auf unser Beispiel: P(König&Karo-König) ≤ P(König)

Ein anderes Beispiel: Angenommen, ich zöge eine Karte aus dem Deck und sagte Ihnen, es sei eine Karo-Karte. Ich biete Ihnen folgende Wahrscheinlichkeiten:

$$\mathrm{C}: \text{Die Karte ist eine Bildkarte.}$$

$$\mathrm{D}: \text{Die Karte ist der Karo} - \text{König.}$$

[2] Sturm T (2009) Selbsttäuschung: Wer ist hier (ir)rational und warum? Studia Philosophica 68, 229–254, S. 243

Kinderleicht: Es gibt 13 Karo-Karten in einem Deck. Es gibt drei Arten von Bildkarten (König, Dame, Bube). Die Wahrscheinlichkeit, dass die gezogene Karte eine Bildkarte ist, vorausgesetzt, Sie wissen, dass es eine Karo-Karte ist, liegt bei 3:13. Es gibt 13 Karo-Karten in einem Deck, das heißt, die Chancen, dass die Karte der Karo-König ist, vorausgesetzt, Sie wissen, dass es eine Karo-Karte ist, liegt bei 1:13, weil es nur einen Karo-König gibt.

Formal sieht das ganze nun so aus:

$$C : P(\text{Bildkarte} \mid \text{Karo}) = 3 : 13$$

$$D : P(\text{König} \mid \text{Karo}) = 1 : 13$$

C hat also die größte Wahrscheinlichkeit.

Fällt Ihnen etwas auf? Unter Anwendung der neuen Information haben Sie Ihre Annahmen oder Gewinnaussichten aktualisiert. Hätten Sie nicht gewusst, dass meine gezogene Karte eine Karo-Karte ist, weil ich es Ihnen vielleicht nicht gesagt habe, hätten Sie Ihre Chancen nur nach den A-priori-Wahrscheinlichkeiten A und B ausgerechnet. Nun aber verfügen Sie über eine neue Information. Sie wissen, dass die Karte eine Karo-Karte ist, und entscheiden sich nun zwischen den aktualisierten Wahrscheinlichkeiten, und das sind die A-posteriori-Wahrscheinlichkeiten C und D.

Was ich Ihnen damit anschaulich vermittelt habe, ist der Satz von Bayes – ein normatives Modell, das nach einer neuen Information oder Beobachtung die Wahrscheinlichkeit für die Wahrheit einer Hypothese aktualisiert.

Wenden wir den Satz von Bayes nun auf unsere Kartenbeispiele C und D an:

Für Beispiel C müssen wir die A-posteriori-Wahrscheinlichkeit von P(Bildkarte|Karo) berechnen. Dafür brauchen wir die A-priori-Wahrscheinlichkeiten von Bildkarte und Karo-Karte. Das ist leicht. Es gibt in einem Deck 12 Bildkarten (König, Dame und Bube, und zwar für jede Farbe, Karo, Herz, Kreuz und Pik). Die A-priori-Wahrscheinlichkeit, eine Bildkarte zu ziehen, liegt also bei 12/52 = 0,23. Es gibt in einem Deck 13 Karo-Karten, also ist die A-priori-Wahr-

scheinlichkeit, eine Karo-Karte zu ziehen, 13/52=0,25. Als nächstes müssen wir die Wahrscheinlichkeit schätzen, welche die bedingte Wahrscheinlichkeit ist, dass die Karte eine Karo-Karte ist, vorausgesetzt, wir wissen, dass die Karte eine Bildkarte ist. Es gibt in einem Deck 12 Bildkarten und drei davon sind Karo-Karten. Also ist P(Karo|Bildkarte) gleich 3/12 und das ergibt 0,25. Nun können wir die A-posteriori-Wahrscheinlichkeit für den Fall C berechnen:

$$\begin{aligned} &P(\text{Bildkarte} \mid \text{Karo}) = P(\text{Karo} \mid \text{Bildkarte})\ [P(\text{Bildkarte})/P(\text{Karo})] \\ &= 0,25(0,23/0,25) \\ &= 0,23 \end{aligned}$$

Die Wahrscheinlichkeiten für C liegen unserer Rechnung zufolge bei 3/13, was das Gleiche ist wie 0,23.

Für Beispiel D brauchen wir die A-priori-Wahrscheinlichkeit, einen König zu ziehen (4/52=0,08), die A-priori-Wahrscheinlichkeit, eine Karo-Karte zu ziehen (0,25), und die Wahrscheinlichkeit, eine Karo-Karte zu ziehen, vorausgesetzt, die Karte ist ein König (1/4=0,25). Die A-posteriori-Wahrscheinlichkeit, einen König zu ziehen, vorausgesetzt, die Karte ist eine Karo-Karte, stellt sich demnach folgendermaßen dar:

$$\begin{aligned} &P(\text{König} / \text{Karo}) = P(\text{Karo} \mid \text{König})\ [P(\text{König})/P(\text{Karo})] \\ &= 0,25(0,08/0,25) \\ &= 0,08 \end{aligned}$$

Die Wahrscheinlichkeiten für D liegen unserer Rechnung zufolge bei 1/13, was das Gleiche ist wie 0,08. Sie sehen, wenn es darum geht, die Häufigkeitsverteilungen der verschiedenen Kartenarten in einem Kartendeck zu schätzen, folgen Sie in Ihren Entscheidungen (intuitiv) dem Satz von Bayes – und das scheint allemal leichter, als jeden einzelnen Gedanken in eine Wahrscheinlichkeit umzurechnen und ihn in eine Bayessche Formel zu packen.

Der Satz von Bayes gilt als das optimale statistische Modell, das als Entscheidungsregel für Risikosituationen angewendet werden kann. Ich will an dieser Stelle darauf hinweisen, dass der Begriff „Risiko" hier nicht zu verwechseln ist mit „Unsicherheit", wie er in der traditionellen Statistik verwendet wird. Bei der Risikoermittlung basieren die Berechnungen auf Wahrscheinlichkeiten, die bekannt sind – wie in unserem Brustkrebsbeispiel. Wenn die Wahrscheinlichkeiten, die mit einem Ereignis oder einer Entscheidung verbunden sind, nicht bekannt sind, lässt sich der Satz von Bayes nicht anwenden. Im Abschnitt 3.2 werden wir „Wahrscheinlichkeiten" der Einfachheit halber als „Häufigkeiten" bezeichnen.

Der Satz von Bayes besagt: Die Wahrscheinlichkeit (Probabilität, P) für das Zutreffen einer Hypothese bei gegebenen Daten zweier Ereignisse A und B (A-posteriori-Wahrscheinlichkeit) ist proportional zum Produkt aus der Wahrscheinlichkeit P(A) und der beobachteten Wahrscheinlichkeit P(B|A) (*likelihood*). Die beobachtete Wahrscheinlichkeit zeigt den Effekt der Daten. Hier die Gleichung:

$$A-posteriori-Wahrscheinlichkeit\,(A|B) = \frac{beobachtete\,Wahrscheinlichkeit\,(B|A) \times A-posteriori\ -Wahrscheinlichkeit\,(A)}{A-priori-Wahrscheinlichkeit\,(B)}$$

Wenden wir nun den Satz von Bayes an, um herauszufinden, wie hoch die Wahrscheinlichkeit ist, dass Sie tatsächlich Brustkrebs haben, unter der Voraussetzung, dass Sie 40 Jahre alt sind und bei Ihrer ersten Mammografie ein positives Ergebnis erhalten haben. Die Prävalenzrate von Brustkrebs in Ihrer Altersgruppe liegt bei 1 %, das heißt,

der positive Vorhersagewert für die Wahrscheinlichkeit von Brustkrebs in Ihrer Altersgruppe liegt bei 0,01. Rund 18 % der Mammografien in dieser Altersgruppe sind positiv, das heißt, der Vorhersagewert, ein positives Testergebnis zu erhalten, liegt bei 0,18. Radiologen stellen in 88 % der Fälle in dieser Altersgruppe eine richtig-positive Krebsdiagnose. Der positive Vorhersagewert für die Wahrscheinlichkeit, von Ihrem Radiologen ein richtig-positives Testergebnis zu bekommen, wenn Sie tatsächlich Brustkrebs haben, liegt also bei 0,88. Jetzt müssen wir diese Zahlen nur noch in unsere Gleichung einsetzen:

$$\begin{aligned} P(\text{Krebs} \mid \text{posTest}) &= P(\text{posTest} \mid \text{Krebs})[P(\text{Krebs})/P(\text{posTest})] \\ &= 0{,}88(0{,}01/0{,}18) \\ &= 0{,}05 \end{aligned}$$

Mit anderen Worten: Im Alter von 40 Jahren liegt die Wahrscheinlichkeit bei positiver Mammografie, tatsächlich an Brustkrebs erkrankt zu sein, bei 5 %. Und angenommen, Sie wären 50 Jahre alt, sähe die Rechnung so aus: $0{,}91(0{,}218/0{,}22) = 0{,}09$ also 9 %.

Beachten Sie, dass die Berechnung immer nur die tatsächliche Prävalenzrate von Brustkrebs in Ihrer Altersgruppe berücksichtigt. Wie die Prävalenzrate (z. B. 2,8 %) zeigt, ist das tatsächliche Vorkommen von Brustkrebs in jeder Altersgruppe extrem selten. Das Lebenszeitrisiko, Krebs zu entwickeln – das heißt, die statistische Wahrscheinlichkeit, im Verlauf des Lebens irgendwann einmal an Krebs zu erkranken –, liegt bei 12 % oder bei 1:8, also bei einem von acht Fällen. Aber nur 1 % der Frauen in der Altersgruppe 40 bis 49 ist tatsächlich an Brustkrebs erkrankt, eine von

112 (1:112). Bis zum Alter von 50 Jahren ist diese Rate nur geringfügig auf etwas mehr als 2 % angestiegen (1:46).

Beim Thema Mammografie streiten sich die Experten: Sollten sich Frauen einer jährlichen Mammografie unterziehen? Und wenn ja, ab welchem Alter sollten sie damit anfangen? „Je früher, desto besser", möchte man aus dem Bauch heraus meinen. Aber es gibt ein Problem: Mammografien sind keine Fotografien. Man kann sie sich nicht einfach ansehen und einen Krebs erkennen. Die Aufnahmen müssen von erfahrenen Radiologen sorgfältig gelesen und interpretiert werden und es gibt eine nicht unerhebliche Fehlerquote durch eine falsche Interpretation. Nach dem Jahresbericht des US-amerikanischen Konsortiums für Brustkrebserkrankungen (Breast Cancer Consortium) von 2009 stellen Radiologen bei 40-jährigen Frauen, die sich erstmals einer Mammografie unterziehen, in rund 8 % der Fälle eine falsch-positive Diagnose. Dies bedeutet, dass rund eine von zwölf dieser Frauen die Diagnose Brustkrebs bekommt, wenn sie tatsächlich gar nicht an Brustkrebs erkrankt ist. Die Falsch-positiv-Rate für 50-jährige Frauen liegt bei 12 % (bei etwa einer von acht Frauen)! Jedes Mal also, wenn Sie sich einer Mammografie unterziehen, gehen Sie das Risiko ein, eine falsch-positive Diagnose zu bekommen. Und Sie erhöhen dieses Risiko bedeutend mit jeder weiteren Mammografie, die Sie im Laufe Ihres Lebens vornehmen lassen. Zum Beispiel hat eine Frau, die sich zehn Mammografien unterzieht, eine 49%ige Wahrscheinlichkeit, mindestens einmal einen falsch-positiven Befund zu bekommen (Christiansen et al. 2000). Wenn Sie im Alter von 40 Jahren mit Mammografien beginnen, bedeutet

das, dass Sie eine beinahe 50%ige Wahrscheinlichkeit haben, mit einer Brustkrebsdiagnose konfrontiert zu werden, was nicht der Fall ist, wenn Sie bereits 50 sind. Aus diesem Grund hat der zuständige Regierungsausschuss in den USA neue Richtlinien herausgegeben, wonach regelmäßige Mammografie-Screenings erst vom 50. und nicht vom 40. Lebensjahr an empfohlen werden, und zwar nicht jährlich, sondern alle zwei Jahre. Diese Empfehlung fand keinen großen Anklang. Frauen entrüsteten sich und ihre Ärzte protestierten. Die Schwierigkeit scheint vor allem darin zu liegen, die Frauen davon zu überzeugen, dass eine positive Mammografie kein eindeutiger Beweis für eine tatsächliche Krebserkrankung ist. (Für weitere Informationen über die richtige Interpretation medizinischer Screening-Ergebnisse siehe Gigerenzer et al. 2008.)

Die Richtlinien zur Früherkennung von Prostatakrebs wurden 2008 ebenfalls geändert. Der Test, der am häufigsten eingesetzt wird, ist der sogenannte PSA-Test. Das prostataspezifische Antigen, abgekürzt PSA, wird von der gesunden wie auch von der kanzerogenen Prostatadrüse produziert. Mit zunehmendem Alter des Mannes treten vermehrt gutartige Prostataerkrankungen und Prostatakrebs auf. Einen Anstieg der PSA-Konzentration in Bezug auf eine gutartige oder kanzerogene Erkrankung zu interpretieren, ist also keine leichte Aufgabe. Vor 2008 wurde empfohlen, den PSA-Test zur Früherkennung von Prostatakrebs ab dem 40. Lebensjahr jährlich vornehmen zu lassen. Doch dann wurden die Ergebnisse der US-Studie „Prostata, Lung, Colorectal and Ovarian Cancer Screening (PLCO)“ veröffentlicht. Im Rahmen dieser Langzeitstudie unter-

suchten die Forscher mehr als 76.000 Männer, die sie nach dem Zufallsprinzip in zwei Gruppen aufteilten: die in der ersten Gruppe erhielten die übliche Behandlung, an denen der zweiten Gruppe wurde sechs Jahre lang einmal jährlich ein PSA-Test und vier Jahre lang einmal jährlich eine digital-rektale Untersuchung durchgeführt. Wie die Forscher nach Auswertung der Folgestudien feststellten, einmal nach sieben und einmal nach zehn Jahren, war der Unterschied in der Gesamtsterberate infolge von Prostatakrebs nicht signifikant.

In seinem Buch *Statistics for management and economics* (2001) zieht Gerald Keller Daten über PSA und Prostatakrebs heran und verwendet diese, um die A-posteriori-Wahrscheinlichkeit zu berechnen, die ein 40- bis 50-jähriger Mann für einen positiven Prostatakrebsbefund hat, wenn sein PSA-Test positiv ist (der Wert bei 4,1 oder höher liegt). Nach Kellers Ergebnis liegt die Wahrscheinlichkeit bei 0,05! Aufgrund von Ergebnissen wie diesem empfiehlt die American Cancer Society (ACS) in ihren aktuellen Richtlinien Folgendes:

> „Männer sollten sich von ihrem Arzt darüber beraten lassen, ob eine Vorsorgeuntersuchung sinnvoll ist, wobei die Entscheidung von bestimmten Risikofaktoren wie z. B. Alter, familiärer Vorbelastung und Rassenzugehörigkeit (Afroamerikaner tragen ein höheres Risiko) abhängt."[2]

Im Oktober 2011 erweiterte die U.S Preventive Services Task Force diese Empfehlungen insoweit, dass die Männer über das Für und Wider der Tests aufgeklärt werden sollten,

[2] http://www.deltastar.nl/pdf/NP/2010/BreastCancer.pdf

um dann für sich selbst entscheiden zu können. Nach Auswertung der Ergebnisse stellte die Task Force fest, dass auch durch regelmäßiges Screening so gut wie keine Verringerung der Todesrate erreicht worden war. Stattdessen werden viel zu viele nichttödliche Tumore entdeckt und aggressiv behandelt, was zu ernsten Nebenwirkungen führt, ausgelöst durch praktisch unnötige Therapien.

Dies bedeutet, dass Arzt und Patient heute eine „bayesische" Entscheidung aufgrund von Wahrscheinlichkeitsinformationen treffen müssen. Doch wie wir gesehen haben, machen selbst Ärzte Fehler, wenn sie ihre Entscheidungen auf probabilistische Informationen stützen. Nach Gerd Gigerenzer liegt dies daran, dass es Informationsdarstellungen in Form von Wahrscheinlichkeiten und Prozenten erst etwa seit dem 18. Jahrhundert gibt. Häufigkeiten lassen sich tatsächlich auch ohne konkrete Zählungen visualisieren (Zacks & Hasher 2002). Das machen wir ständig und ganz automatisch – ein Vorgang, den die Fachwelt mit dem englischen Begriff *subitizing* bezeichnet. Wenn Sie sich beispielsweise in Ihrer unmittelbaren Umgebung umschauen, sehen Sie vielleicht vier Leute, die ein rotes T-Shirt tragen, und zwei mit einem blauen T-Shirt. Sie sehen nicht 66 % in einem roten T-Shirt und 24 % in einem blauen. Sie sehen keine Wahrscheinlichkeit von 0,6 für rote T-Shirts und von 0,24 für blaue. Natürlich lässt sich die Häufigkeitsverteilung mathematisch in Prozent darstellen, aber was Ihre Wahrnehmung registriert, ihr kognitives System, das Informationen von Reizen aus der Umwelt gewinnt und verarbeitet, sind Mengen oder Häufigkeiten. Und es „versteht" diese Häufigkeiten.

Tab. 3.2 Brustkrebsprävalenz nach Alter der Patienten.

Alter	Prävalenz	
30–39	1,7 von 1000	1 von 588
40–49	8,9 von 1000	1 von 112
50–59	21,8 von 1000	1 von 46
60–69	37,5 von 1000	1 von 26

Tabelle 3.2 zeigt die Darstellung der Brustkrebsdaten, wenn wir Prävalenz in Häufigkeiten (oder *odds*) ausdrücken.

„Bayesisch" denken lernen

Was wäre nun, wenn wir die Wahrscheinlichkeitsinformationen aus der Tab. 3.2 in Häufigkeitsinformationen übersetzen wollten, um herauszufinden, was es mit einer positiven Mammografie auf sich hat? Wären wir dann in der Lage, klügere Entscheidungen zu treffen? Versuchen wir es einmal:

> Stellen wir uns 1000 Frauen vor, die sich mit 40 erstmals einer Mammografie zur Brustkrebsvorsorge unterziehen.
> Neun von ihnen haben Brustkrebs und erhalten einen positiven Mammografiebefund.
> 178 von ihnen haben keinen Brustkrebs und bekommen einen positiven Mammografiebefund.
> Wie hoch liegt die Wahrscheinlichkeit, dass eine Frau mit einer positiven Mammografie auch tatsächlich Brustkrebs hat?

Ganz leicht: Die Antwort lautet 9 von 178 (was 5 % entspricht, oder 1 von 20). Das heißt im Klartext: Die Wahrscheinlichkeit, dass eine Frau mit einer positiven Mammografie tatsächlich Brustkrebs hat, liegt bei 1:20, und bei 19:20, dass Sie nicht daran erkrankt sind. Sobald man entscheidungsrelevante Informationen in einem Häufigkeitsformat wie diesem präsentiert, steigt die Trefferquote in der Einschätzung der Wahrscheinlichkeiten rasant an, und das ist bei Ärzten nicht anders (Chase et al. 1998; Gigerenzer & Hoffrage 1995). Darstellungen in diesem Format, so betonen die Autoren, erleichtern die Aussagen ungemein, da sie uns weniger anfällig für fehlerhafte Risikoeinschätzungen machen.

Zurück zu der 9 und der 178 und dazu, woher diese beiden Zahlen kommen: Wir wissen, dass die Brustkrebsrate für diese Altersgruppe bei 1 % liegt, und das bedeutet, dass 10 Frauen von 1000 voraussichtlich an Brustkrebs erkranken werden, und 990 nicht. Radiologen stellen in dieser Altersgruppe aber in 88 % der Fälle eine richtig-positive Diagnose ($10 \times 0{,}88 = 9$). Neun dieser Frauen werden also an Brustkrebs erkranken und eine richtig-positive Diagnose bekommen, eine aber wird eine falsch-negative Diagnose erhalten. Das heißt, sie ist tatsächlich an Brustkrebs erkrankt, der Befund wird aber negativ interpdretiert. Wir wissen außerdem, dass Radiologen in 18 % der Fälle dieser Altersgruppe fälschlicherweise Brustkrebs diagnostizieren, wenn er gar nicht vorhanden ist. Von den 990 brustkrebsfreien Frauen bekommen also 178 einen positiven Mammografiebefund ($990 \times 0{,}18 = 178$). Diese 178 bekommen im Anschluss möglicherweise jede Menge unnötiger Therapien, ganz zu schweigen von einem großen Schock über die

Diagnose. Spielen wir nun die gleiche Situation für Frauen im Alter von 50 durch:

> Stellen wir uns 1000 Frauen vor, die sich mit 50 erstmals einer Mammografie zur Brustkrebsvorsorge unterziehen.
> 20 von ihnen haben Brustkrebs und erhalten einen positiven Mammografiebefund.
> 215 von ihnen haben keinen Brustkrebs und bekommen einen positiven Mammografiebefund.
> Wie hoch liegt die Wahrscheinlichkeit, dass eine Frau mit einer positiven Mammografie auch tatsächlich Brustkrebs hat?

Ganz leicht: Die Antwort lautet 20 von 235 der positiv getesteten Frauen (was rund 9 % entspricht, oder 1 von 11). Man beachte auch hier, dass 215 Frauen ein falsch-positives Ergebnis und unnötige Therapien bekommen werden. Aber 1 von 11 ist besorgniserregender als 1 von 20. Und eben weil das Erkrankungsrisiko mit dem Alter zunimmt, empfiehlt sich die Mammografie insbesondere für Frauen ab 50 Jahren.

Ich will Ihnen noch weitere Beispiele geben, die zeigen, wie drastisch sich unsere Fähigkeit, „bayesisch" zu denken, verbessert, wenn uns Informationen im Häufigkeitsformat vorliegen. Auch die folgenden Beispiele sind Gigerenzers Buch *Calculated risks* entnommen:

> Frage: Haben Männer mit einem zu hohen Cholesterinspiegel Grund zur Panik, wenn man weiß, dass ein zu hoher Cholesterinspiegel das Risiko, einen Herzinfarkt zu erleiden, um 50 % erhöht?

Antwort: 6 von 100 Männern mit einem erhöhten Cholesterinspiegel werden innerhalb von 10 Jahren einen Herzinfarkt erleiden gegenüber 4 von 100 mit einem normalen Cholesterinspiegel. In absoluten Zahlen heißt dies: Das erhöhte Risiko liegt bei nur 2 von 100 – oder 2 %. Man kann es auch von der anderen Seite betrachten: Von den 100 Männern mit einem hohen Cholesterinspiegel werden 94 % keinen Herzinfarkt erleiden.

Frage: HIV-Tests sind zu 99,9 % genau. Ihr HIV-Test fällt positiv aus, obwohl Sie keinerlei bekannte Risikofaktoren haben. Wie hoch ist die Wahrscheinlichkeit, dass Sie AIDS haben, wenn 0,01 % der Menschen mit keinerlei bekannten Risikofaktoren infiziert sind?

Antwort: 50:50. Nehmen wir 10.000 Menschen mit keinerlei bekannten Risikofaktoren. Einer dieser Menschen hat AIDS; er wird so gut wie sicher ein positives Testergebnis erhalten. Von den verbleibenden 9999 Menschen wird einer ebenfalls positiv getestet werden. Folglich ist die Wahrscheinlichkeit, eine richtig-positive AIDS-Diagnose zu bekommen, 1:2. Ein positiver AIDS-Test, obgleich begründeter Anlass zur Besorgnis, ist bei Weitem kein Todesurteil.

Frage: Im Strafprozess gegen den ehemaligen Footballspieler O.J. Simpson, dem vorgeworfen wurde, seine Ehefrau ermordet zu haben, gelang es dem Verteidiger Alan Dershowitz, die Geschworenen davon überzeugen, dass die ermittelten Beweise irrelevant seien, denn auch bei 2,5 bis 4 Millionen bekannter Fälle von häuslicher Gewalt, so Dershowitz, gebe es nur 1432 Morde. Und damit, so führte er im Weiteren aus, „ist die Prozentzahl der Männer, die ihre Frauen schlagen und misshandeln und sie dann auch ermorden, verschwindend gering – und liegt höchstwahrscheinlich bei unter 1 von 2500." Hatte er Recht?

Antwort: Nehmen wir 100.000 misshandelte Frauen. Vierzig von ihnen werden im Laufe eines Jahres von ihrem Ehemann/Lebensgefährten ermordet. Fünf von ihnen werden von einer anderen Person ermordet. Damit werden 40 bzw. 45 der ermordeten und misshandelten Frauen von ihren Peinigern getötet, und in nur 1 von 9 Fällen ist der Mörder eine andere Person als der Peiniger.

Wenn wir nicht „bayesisch" denken

Damit es noch klarer wird, schauen wir uns anhand einer Studie von Kahneman und Tversky (1973) einmal an, was wir tun, wenn wir Entscheidungen auf Basis von Wahrscheinlichkeitsinformationen treffen sollen. Versuchen Sie gleich selbst, die folgenden Fragen zu beantworten:

Eine Gruppe von Psychologen führte mit 30 Ingenieuren und 70 Anwälten, allesamt in ihrem jeweiligen Fachgebiet sehr erfolgreich, Persönlichkeitstests durch. Auf Grundlage der gegebenen Informationen verfassten sie eine kurze Beschreibung für jeden der 30 Ingenieure und 70 Anwälte. Ich lege Ihnen nun fünf Beschreibungen von Personen vor, die zufällig aus eben dieser Gruppe von 100 Personen gezogen wurden, und möchte Sie bitten, auf einer Skala von 0 bis 100 jeweils anzugeben, wie hoch Sie die Wahrscheinlichkeit einschätzen, dass die beschriebene Person ein Ingenieur ist.
Kahneman und Tversky legten diese Beschreibungen bei ansonsten gleicher Aufgabenstellung auch einer Gruppe von Experten vor, die eine sehr hohe Trefferquote erzielten. Sie bekommen also einen zusätzlichen Punkt, wenn

Sie mit Ihren Einschätzungen nahe an die der Experten herankommen.

1. Jack ist ein Mann von 45 Jahren. Er ist an politischen und sozialen Fragen nicht interessiert und verbringt seine Freizeit mit seinen zahlreichen Hobbys, zu denen Heimwerken, Segeln und Denksportaufgaben gehören. Die Wahrscheinlichkeit, dass Jack einer der 30 Ingenieure aus der Gruppe der 100 Personen ist, liegt bei _____%.
2. Tom ist 34 Jahre alt. Er ist intelligent, aber fantasielos, eigenbrötlerisch und insgesamt recht farblos. In der Schule war er gut in Mathe, aber schlecht in Sozialkunde und Geisteswissenschaften. Die Wahrscheinlichkeit, dass Tom einer der 30 Ingenieure aus der Gruppe der 100 Personen ist, liegt bei _____ %.
3. Nun nehmen wir an, Sie haben keinerlei Information über irgendeine dieser zufällig ausgewählten Personen. Die Wahrscheinlichkeit, dass diese Person eine der 30 Ingenieure aus der Gruppe der 100 Personen ist, liegt bei _____ %.

Wenn Sie wie die Mehrheit der Teilnehmer dieser Studie geantwortet haben, haben Sie die Wahrscheinlichkeit für Jack unter 30 % geschätzt, für Tom über 30 % und für die dritte Aussage 50 %. Die Ausgangssituation aber gibt 30 Ingenieure und 70 Rechtsanwälte vor. Die Wahrscheinlichkeit, dass irgendeine dieser 100 Personen Ingenieur ist, liegt also ohnehin bei 30 %. Tversky und Kahneman zeigten, dass wir in Entscheidungssituationen wie dieser die A-priori-Wahrscheinlichkeiten (die Grundraten für die unterschiedlichen Fälle) tendenziell ignorieren. Mal angenommen, es wäre um ein Kartendeck gegangen und man hätte die Teilnehmer gebeten zu schätzen, mit welcher Wahrscheinlichkeit sie eine Karo-Karte ziehen würden.

Wenn sie so geantwortet hätten wie in der Studie, hätten sie die Wahrscheinlichkeit auf eine Karo-Karte in Fall 1 auf weniger als 25 %, in Fall 2 auf mehr als 25 % und für Fall 3 auf 50 % geschätzt.

Aber wir brauchen diese Zahleninformationen nicht einmal heranzuziehen, um Abweichungen von der „bayesischen" Logik zu demonstrieren. Betrachten wir die Konjunktionsregel, die besagt, dass wir das gemeinsame Eintreten zweier (unabhängiger) Ereignisse für wahrscheinlicher halten als das Eintreten eines der beiden Einzelereignisse. Im Rahmen einer Studienreihe bekamen die Teilnehmer eine Liste mit Aussagen vorgelegt, die sie in der Reihenfolge ihrer Wahrscheinlichkeit ordnen sollten, die wahrscheinlichste sollte ganz oben stehen, die unwahrscheinlichste am Ende. Hier ein Beispiel (Tversky & Kahneman 1983):

> Tom ist 34 Jahre alt. Er ist intelligent, aber fantasielos, eigenbrötlerisch und insgesamt recht farblos. In der Schule war er gut in Mathe, aber schlecht in Sozialkunde und Geisteswissenschaften.
> A: Tom ist Buchhalter.
> B: Tom ist Arzt, der in seiner Freizeit Poker spielt.
> C: Tom spielt in seiner Freizeit Jazzmusik.
> D: Tom ist Architekt.
> E: Tom ist Buchhalter, der in seiner Freizeit Jazzmusik spielt.
> F: Toms Hobby ist Bergsteigen.

Die große Mehrheit der Teilnehmer gab an, dass sie Tom am wahrscheinlichsten für einen Buchhalter hält, der in seiner Freizeit Jazzmusik spielt, und nicht nur für einen

Hobby-Jazzmusiker. Das heißt, sie hielten die Konjunktion (Verbindung) „Buchhalter und Jazzmusiker" für wahrscheinlicher als „Hobby-Jazzmusiker". Diese Folgerung nennt man Konjunktionsfehler – verbundene Ereignisse werden als wahrscheinlicher angesehen als eines von ihnen. Wir können dies symbolisieren als $P(A\&J) > P(J)$.

Noch ein Beispiel:

> Linda ist 31 Jahre alt, Single, offen und sehr intelligent. Sie hat einen Masterabschluss in Philosophie. Als Studentin war sie sehr engagiert bei Themen wie Diskriminierung und sozialer Gerechtigkeit und nahm auch an Anti-Atomkraft-Demonstrationen teil.
> Ordnen Sie die folgenden Aussagen nach ihrer Wahrscheinlichkeit, beginnen Sie mit der wahrscheinlichsten und schließen Sie mit der unwahrscheinlichsten.
> A: Linda ist Grundschullehrerin.
> B: Linda arbeitet in einer Buchhandlung und gibt Yoga-Kurse.
> C: Linda ist aktiv in der Frauenbewegung.
> D: Linda ist Heilpädagogin.
> E: Linda ist Mitglied in einer Frauenrechtsorganisation.
> F: Linda ist Bankkassiererin.
> G: Linda ist Versicherungsmaklerin.
> H: Linda ist Bankkassiererin und aktiv in der Frauenbewegung.

Tversky und Kahneman (1983) führten diese Studie mit Erstsemestern an der University of British Columbia durch: 92 % stuften H höher ein als F. Die gleiche Studie führten sie auch mit Doktoranden im Fach Entscheidungswissenschaften an der Stanford University Business School durch,

von denen alle bereits Hauptseminare in Wahrscheinlichkeitsberechnung und Statistik belegt hatten. Doch das war ganz egal; auch sie erlagen dem Konjunktionsfehler und stuften mit 83 % H ebenfalls höher ein als F.

Warum versagen wir so kläglich darin, diese Aufgaben zu lösen? Aus drei Gründen. Zum Ersten verlangen diese Aufgaben von uns, Schlussfolgerungen aufgrund von Wahrscheinlichkeiten und nicht aufgrund von Häufigkeiten zu treffen. Und wie wir gesehen haben, tun sich die meisten Menschen schwer damit, in Wahrscheinlichkeiten zu denken. Zum Zweiten geht aus den Beschreibungen nicht eindeutig hervor, dass die Karten nach dem Zufallsprinzip gezogen werden. Sobald dies jedoch ausdrücklich formuliert ist, lenken die Teilnehmer ihr Augenmerk verstärkt auf die Basisraten und schneiden besser ab (Gigerenzer et al. 1988). Und zum Dritten, und das ist vielleicht am wichtigsten, stehen die Wahrscheinlichkeitsinformationen hier den subjektiven Ähnlichkeitsurteilen entgegen. Wir verstehen es hervorragend, nach Ähnlichkeiten zu suchen und Dinge nach der Ähnlichkeit mit einem Prototyp zu klassifizieren. Das tun wir im Grunde ganz automatisch. Obwohl die Probanden der Studie aufgefordert waren, ein Wahrscheinlichkeitsurteil zu treffen, lädt die Formulierung der Aufgabenstellung sie implizit dazu ein, aufgrund von Klassifizierungen ein Ähnlichkeitsurteil zu fällen.

Tversky und Kahneman bezeichnen diese Art der Urteilsfindung als Repräsentativitätsheuristik: Eine Person beurteilt die Wahrscheinlichkeit eines Ereignisses nach dem Ausmaß, in dem es in wesentlichen Eigenschaften seiner Grundgesamtheit ähnlich ist. Fallzahlen und A-priori-Wahrscheinlichkeiten werden dabei völlig ignoriert. Im ers-

ten Beispiel beginnen wir ganz automatisch, die beschriebenen Eigenschaften mit denen eines typischen Ingenieurs zu vergleichen und antworten entsprechend der Ähnlichkeiten, die wir bei einem typischen Ingenieur finden. Im zweiten Beispiel antworten wir entsprechend der Ähnlichkeiten, die wir bei einem typischen Mathematiker finden. Und das dritte Beispiel animiert uns zu überlegen, welche typischen Eigenschaften ein Bankkassierer haben könnte.

Was passiert, wenn die Probanden nun kein Wahrscheinlichkeitsurteil, sondern eine Häufigkeitsschätzung abgeben sollen – wenn sie ihre Schlussfolgerungen also nicht aufgrund von Wahrscheinlichkeiten, sondern aufgrund von Häufigkeiten ziehen sollen? Hertwig und Gigerenzer (1999) legten ihren Versuchspersonen eine Aufgabe vor, in der 200 Frauen beschrieben wurden, die alle auf die Beschreibung von Linda passten. Die Fragen dazu lauteten: Wie viele von den 200 Frauen sind Bankkassiererinnen? Wie viele von den 200 Frauen sind in der Frauenbewegung aktiv? Und wie viele von den 200 Frauen sind Bankkassiererinnen und in der Frauenbewegung aktiv? Keiner der Probanden verletzte die Konjunktionsregel, wenn die Fragen in Form von natürlichen Häufigkeiten präsentiert wurden. Der springende Punkt ist, dass Schlussfolgerungen auf der Basis von Häufigkeiten offenbar besser gelingen als aufgrund von Wahrscheinlichkeiten. In einer Folgestudie stellten Sloman et al. (2003) fest, dass 70 % der Versuchspersonen die Konjunktionsregel verletzten, wenn sie aufgefordert waren, Häufigkeitsschätzungen nach einer Präferenzordnung aufgrund von bekannten Zahlen (Basisraten) vorzunehmen, aber nur knapp über 30 %, wenn sie

Wahrscheinlichkeiten freiweg schätzen sollten (wie in den Studien von Hertwig und Gigerenzer).

Die Formulierung der Frage bestimmt, ob wir richtige oder falsche Schlüsse ziehen

„Linda" und „Tom" zeigen uns, wie leicht kleine Änderungen in einer ansonsten inhaltsgleichen Formulierung unsere Urteile beeinflussen können. Sie führen uns den sogenannten Framing-Effekt (oder Rahmungseffekt) vor Augen: Unsere Entscheidungen werden sehr viel stärker davon beeinflusst, wie ein Problem formuliert (beschrieben) ist, als von den objektiven Daten, die darin enthalten sind. Die stärksten Framing-Effekte treten meist dann auf, wenn die als real präsentierten Wahrscheinlichkeiten die tief verankerten subjektiven Wahrscheinlichkeiten in unserer kognitiven Architektur verzerren. Das beste Beispiel für solche kognitiven Verzerrungen ist die Verlustaversion, die dafür sorgt, dass viele Menschen übermäßig darauf bedacht sind, Verluste zu vermeiden, um Gewinne zu erzielen (Tversky & Kahneman 1981). Einige Studien legen nahe, dass Verluste, psychologisch gesehen, doppelt so stark empfunden werden wie Gewinne.

Hier ein deutliches Beispiel dafür:

> Stellen Sie sich vor, Sie sind Lungenkrebspatient. Welche der folgenden beiden Möglichkeiten würden Sie vorziehen?

> A Operation: Von 100 Personen, die operiert werden, überleben 90 die Operation. 68 überleben das erste Jahr. 34 überleben die nächsten fünf Jahre.
> B Strahlentherapie: Von 100 Patienten, die bestrahlt werden, überleben alle die Behandlung. 77 überleben das erste Jahr. 22 überleben die nächsten fünf Jahre.

44 % der Befragten bevorzugten die Strahlentherapie gegenüber der Operation. Lesen Sie nun folgende Beschreibung:

> Stellen Sie sich vor, Sie sind Lungenkrebspatient. Welche der folgenden beiden Möglichkeiten würden Sie vorziehen?
> A Operation: Von 100 Patienten, die operiert werden, sterben 10 während der Operation. 32 Patienten sterben innerhalb eines Jahres. 66 Patienten sterben innerhalb der nächsten fünf Jahre.
> B Strahlentherapie: Von 100 Patienten, die bestrahlt werden, sterben 0 während der Behandlung. 23 sterben innerhalb eines Jahres. 78 Patienten sterben innerhalb der nächsten fünf Jahre.

Bei dieser Beschreibung favorisierten nur 15 % die Strahlentherapie gegenüber der Operation. Und es gab diesbezüglich keinen Unterschied zwischen Patienten oder Ärzten. Aber schauen wir uns die Zahlen einmal genauer an. Es sind genau die gleichen: „90 % überleben" ist dasselbe wie „10 sterben". „68 % überleben" ist dasselbe wie „32 sterben". Und so weiter. Das Problem ist genau dasselbe, nur einmal so und einmal so dargestellt – einmal mit dem Fokus auf den Überlebensraten und einmal mit dem Fokus auf den Sterberaten.

Diese Art der Verlustaversion beeinflusst auch sehr stark unseren Umgang mit Geld. Betrachten wir folgende Situation:

> Sie haben soeben 1000 Dollar erhalten und können nun zwischen zwei Optionen wählen:
> Sie gehen auf Risiko:
> Kopf = Sie erhalten zusätzliche 1000 Dollar
> Zahl = Sie erhalten kein zusätzliches Geld
> Sie gehen auf Nummer sicher:
> Sie gewinnen garantiert zusätzliche 500 Dollar

Die meisten Menschen entscheiden sich für die sichere Nummer – Dollar nehmen und damit gut!

Betrachten wir nun diese Situation:

> Sie haben soeben 2000 Dollar erhalten und können nun zwischen zwei Optionen wählen:
> Sie gehen auf Risiko:
> Kopf = Sie verlieren 1000 Dollar
> Zahl = Sie verlieren nichts
> Sie gehen auf Nummer sicher:
> Sie verlieren garantiert 500 Dollar

Die meisten Menschen entscheiden sich hier für das Risikospiel, denn sie wollen es vermeiden, 500 Dollar zu verlieren. Aber genau da liegt das Problem: Es ist exakt das gleiche Spiel. In der Risikovariante stehen die Chancen in beiden Situationen 50:50, mit 1000 Dollar bzw. 2000 Dollar abzuschließen, in der sicheren Variante schließt man mit 1500 Dollar ab.

Einfacher ausgedrückt: Wenn es um mögliche Verluste geht, wählen wir das Risiko (um einen vermeintlichen Verlust zu vermeiden). Im obigen Beispiel wählen die meisten Menschen die Risikovariante und nehmen damit in Kauf, das Doppelte des sicheren Verlusts zu verlieren, denn sie hoffen, dass die Münze mit Zahl nach oben fällt. Die unterschiedliche Wahrnehmung von Gewinnen und Verlusten zeigt sich auch im Anlageverhalten, denn sie bedeutet, dass wir Geldeinlagen oder Immobilien halten, solange sie Verluste bringen in der Hoffnung, dass sie sich regenerieren und wieder ihren ursprünglichen Wert erhalten. Doch damit beschleunigen wir den Verfall der Kurse immer mehr, wie wir in der schweren Wirtschaftskrise 2008 erfahren haben.

Wir Menschen sind aber nicht die einzigen, die sich so verhalten. Nach Studien der Verhaltensforscherin Laurie Santos machen es Kapuzineraffen ganz genauso. Santos und ihre Kollegen konzipierten eine Studienreihe mit verschiedenen Experimenten. In einem dieser Experimente sollten sich Kapuzineraffen zwischen zwei Angeboten entscheiden. Das Angebot kam in Form von Weintrauben, die jeweils von zwei verschiedenen Forscherkollegen dargereicht wurden (Lakshminarayanan et al. 2011). Gewinnszenario: Der eine Forscher legte immer eine Traube dazu, bevor er dann beide Trauben abgab. Ging der Affe auf das Angebot ein, bekam er also immer zwei Trauben (sicherer Gewinn). Der andere Forscher gab manchmal keine weitere Traube dazu, manchmal aber auch zwei. Ging der Affe auf dieses Angebot ein, bekam er manchmal drei Trauben und manchmal eben nur eine (unsicherer Gewinn). Verlustszenario: Die Forscher offerierten den Affen jeweils drei Trauben. Der

eine Forscher nahm immer eine Traube weg, bevor er dem Affen die Trauben gab. Ging der Affe auf das Angebot ein, bekam er am Ende also immer zwei Trauben (sicherer Verlust). Der andere Forscher nahm manchmal zwei Trauben weg und manchmal keine. Ging der Affe auf das Angebot ein, bekam er manchmal drei Trauben und manchmal nur eine (unsicherer Verlust).

Und was kam dabei heraus? Überraschenderweise verhielten sich die Affen in diesem Experiment genau wie wir Menschen: Sie entschieden sich für den sicheren Gewinn im Gewinnszenario und für den unsicheren Verlust im Verlustszenario. Dies bedeutet, so stellt Santos heraus, dass diese Strategie seit mindestens 35 Millionen Jahren existiert, denn so lange ist es her, dass Mensch und Kapuzineraffe einen gemeinsamen Vorfahren im evolutionären Stammbaum hatten.

Kahneman und Tversky (1979) demonstrierten die mathematischen Zusammenhänge dieser kognitiven Verzerrungen in Situationen der Unsicherheit in der sogenannten Neuen Erwartungstheorie (*prospect theory*), einem hervorragenden wissenschaftlichen Modell, für das Kahneman 2002 mit dem Nobelpreis für Wirtschaftswissenschaften geehrt wurde (Tversky war bereits verstorben). In dieser Theorie gibt es zwei Schlüsselmechanismen: 1) Wir empfinden Verluste stärker als vergleichbare Gewinne (10 Dollar gewinnen, fühlt sich gut an, aber 10 Dollar verlieren, fühlt sich wesentlich schlimmer an). 2) Wir reagieren auf Veränderungen mit relativen Gewinnen oder Verlusten, nicht mit absoluten (10 Dollar verlieren oder gewinnen wiegt mehr, wenn wir nur 20 Dollar haben, aber weniger, wenn wir 200 Dollar haben). Das bedeutet: Wir finden zu einer Entschei-

dung, indem wir die verschiedenen Alternativen nicht absolut, sondern relativ zu einem bestimmten Referenzpunkt bewerten.

Kahneman (2003) postulierte zudem eine Zwei-Prozess-Theorie (*dual-process theory*), um einige der Besonderheiten im menschlichen Entscheidungsprozess zu erklären. Kernaussage der Theorie ist, dass es zwei unterschiedliche Denksysteme gibt, die uns in unseren Entscheidungen steuern – ein schnelles (System 1) und ein langsameres (System 2). Laut Kahneman erfolgen die Entscheidungen des schnellen Systems auf der Grundlage von Emotionen, Intuitionen und Heuristiken (wie etwa auf kausalen Interpretationen von Ereignissen und Bildern). Die Entscheidungen des langsameren Systems erfolgen auf der Grundlage der abstrakten, insbesondere auf dessen logischer oder statistischer, Natur.

Wenn System 1 eingreift, kann es sein, dass wir dem eigenen Urteil übermäßig vertrauen. *Overconfidence bias* heißt dieses Phänomen: Ob Laie oder Experte, es liegt offenbar in der Natur des Menschen, die eigenen Fähigkeiten und Urteile stark zu überschätzen. Das führt häufig dazu, dass wir die Suche nach Antworten einstellen, noch bevor wir alle verfügbaren Erkenntnisse zusammentragen können. Hier einige Beispiele:

> Welche Stadt in der jeweiligen Gruppe A und B hat mehr Einwohner? Geben Sie auf einer Skala von 1 bis 10 an, wie sicher Sie sich in Ihrem Urteil sind: 10 = absolut sicher, 0 = geraten.
> A: Las Vegas, Sydney, Berlin
> B: Miami, Melbourne, Heidelberg

Welches historische Ereignis in der jeweiligen Gruppe A und B geschah zuerst? Geben Sie auf einer Skala von 1 bis 10 an, wie sicher Sie sich in Ihrem Urteil sind: 10 = absolut sicher, 0 = geraten.

A: Unterzeichnung der Magna Carta, Tod Napoleons, Ermordung Lincolns

B: Geburt des Propheten Mohammed, Kauf von Louisiana, Geburt von Königin Victoria

Tabelle 3.3 zeigt die Ergebnisse in Bezug auf die subjektive Antwortsicherheit und die tatsächliche Antwortrichtigkeit.

Wie man daran deutlich sieht, liegt die Sicherheit bezüglich der Richtigkeit der Antworten durchweg höher, als die tatsächliche Häufigkeit der richtigen Antworten. Das Ergebnis änderte sich auch dann nicht, als die Probanden etwa den Hinweis erhielten, dass wir alle gemeinhin dazu neigen, das eigene Urteil zu überschätzen, oder wenn man ihnen Geld als Belohnung für die richtige Antwort bot. Dieses Phänomen wurde in zahlreichen Studien mit unterschiedlichen Personengruppen nachgewiesen, darunter Erstsemester, Doktoranden, Ärzte und sogar CIA-Analysten (s. Lichtenstein et al. 1982).

Dabei kommt es aber auch weitgehend darauf an, welche Fragen man stellt. Juslin et al. (2000) analysierten 135 Studien zu diesem Phänomen und fanden heraus, dass bestimmte Frageelemente so ausgesucht wurden, dass sie überrepräsentiert waren und mit hoher Wahrscheinlichkeit zu Falschantworten führten. Betrachten wir noch einmal das erste Beispiel mit der Frage nach der Bevölkerungsgröße der Städte. Eine faire Methode, das Allgemeinwissen und die Antwortsicherheit der Probanden in diesem Bereich zu prüfen, könnte darin bestehen, zuerst sämtliche deutsche

Tab. 3.3 Verhältnis zwischen Antwortsicherheit und Antwortrichtigkeit.

Sicherheit (%)	Richtigkeit (%)
100	80
90	70
80	60

Städte mit mehr als 100.000 Einwohner in einem Datenpool zu sammeln, um daraus dann ein Prüfmuster zu generieren, indem man die einzelnen Städte per Zufallsauswahl aus diesem Datenpool herausfiltert. Werden die einzelnen Elemente einer Frage auf diese Weise ausgesucht, verschwindet der *overconfidence*-Effekt zumeist.

Warum aber scheinen wir diesem Effekt offenbar aufzusitzen und uns der Richtigkeit unserer Antworten fälschlicherweise so überaus sicher zu sein? Vielleicht liegt es daran, dass sich die heuristischen Methoden, die wir anwenden, in unserem alltäglichen Umfeld als besonders adaptiv erweisen und uns für gewöhnlich recht gute Lösungen liefern. Gerd Gigerenzer, Direktor der Abteilung „Adaptives Verhalten und Kognition" am Berliner Max-Planck-Institut für Bildungsforschung, konnte im Rahmen eines Forschungsprogramms eine Reihe von schnellen und sparsamen Heuristiken zu identifizieren – schnell, weil sie das Problem innerhalb weniger Sekunden lösen, und sparsam, weil sie wenig Information benötigen –, die ebenso effektiv sind wie heuristische Algorithmen, die (in einem seriellen Versuchsaufbau) sehr viel mehr Information und Zeit brauchen (Gigerenzer et al. 2000). Der entscheidende Punkt ist, dass diese Heuristiken beobachtete Regelmäßigkeiten nutzen, um intelligente Schlüsse zu ziehen.

Zu den wesentlichen Merkmalen dieser heuristischen Methoden gehören eine begrenzte, sprich auf die Umwelt beschränkte, Such- und eine Stoppregel. Ebenso wie physiologische Systeme darauf angewiesen sind, *Energie* effizient umzusetzen, sind mentale Systeme darauf angewiesen, *Informationen* effizient zu verarbeiten. Gewinnung und Verarbeitung von Information verursachen Kosten, richtige Urteile und Entscheidungen aber bringen Nutzen. Im realen Leben erfolgt ein ständiges Gewichten zwischen beidem. Die Stoppregel besteht schlicht und einfach darin, dass wir die Suche nach Informationen einstellen, sobald die Kosten dafür so groß sind wie der Nutzen daraus, oder wenn die Kosten den aus der weiteren Suche entstehenden Nutzen möglicherweise übersteigen. Ein endloses Grübeln und Gewichten nämlich kann den Entscheidungsprozess ausbremsen und es droht „Paralyse durch Analyse", wie man so schön sagt.

In Kapitel 5 seines 2000 erschienen Buches *Simple heuristics that make us smart* beschreiben Gerd Gigerenzer und Peter M. Todd die Ergebnisse einer Studie, in der unterschiedliche Entscheidungsmodelle zum Einsatz kamen, um empirische Vorhersagen über das Entscheidungsverhalten zu treffen: Es galt, im Rahmen einer Vergleichsaufgabe eines von zwei Elementen, die zufällig aus 20 Datensätzen ausgewählt wurden, höher zu bewerten als das andere. Die Vergleichsaufgaben waren als Frage formuliert, zum Beispiel:

Welche Stadt hat mehr Einwohner?
Welche Schule hat mehr Schulabbrecher?
Welche Autobahn verzeichnet höhere Unfallraten?
Welche Person ist attraktiver?

Datensätze	
Schulabbrecherquoten	Fischfertilität
Obdachlosenzahlen	Schlaf (Säugetiere)
Sterberaten	Biogas durch Kühe
Einwohnerzahlen (Städte)	Artenvielfalt
Immobilienpreise	Niederschlagsmengen
Grundstückspreise	Schadstoffmengen (Los Angeles)
Professorengehälter	Ozonmengen (San Francisco)
Attraktivität (Männer)	Autounfälle
Attraktivität (Frauen)	Benzinverbrauch
Autounfälle	Fettleibigkeit bei 18-Jährigen
Benzinverbrauch	Körperfett

Die Entscheidungsmodelle umfassten zwei normative Strategien und zwei heuristische Strategien.

- normative Strategien:
  - multiple Regression: ein statistisches Verfahren, das der Vorhersage der Werte einer (abhängigen) Variablen aus mehreren fest vorgegebenen Covariablen dient, unter Verwendung der „Methode der kleinsten Quadrate“ zur Schätzung der Parameter (auch als Kleinste-Quadrate-Schätzung bezeichnet)
  - Dawes-Regel: eine vereinfachte Regression, die *unit-weights* nutzt – das ungewichtete Abzählen der Hinweise (+1 oder –1)–, anstatt nach *optimal weights* zu suchen – nach einer optimalen Suchreihenfolge oder Ordnung in einem Datensatz. (Im Grunde genommen, werden positive Hinweise addiert und negative Hinweise subtrahiert.)

- heuristische Strategien:
  - *take the best, forget the rest*: eine Strategie, die Kriterien (*cues*) auf ihre Brauchbarkeit abklopft, und eingestellt wird, sobald ein Kriterium gefunden ist, das eines von zwei Elementen aus einem Datensatz diskriminiert
  - minimalistische Heuristik: eine Strategie, die in einer zufälligen Reihenfolge nach Kriterien sucht und eingestellt wird, sobald ein Kriterium gefunden ist, das eines von zwei Elementen aus einem Datensatz diskriminiert.

Die Ergebnisse waren erstaunlich: Die heuristischen Modelle waren etwas genauer als die sehr viel zeit- und verarbeitungsintensiveren normativen Modelle. Die Vorhersageraten über die Akkuratheit einer Entscheidungsstrategie in einer Entscheidungsaufgabe lagen für *take the best* bei 71 %, für die *minimalistische Heuristik* bei 65 %. Im Gegensatz dazu lagen die Vorhersageraten über die Akkuratheit für die *multiple Regression* bei 68 % und für *Dawes*-Regel bei 69 %. Diese Ergebnisse zeigen sehr deutlich, dass heuristische Entscheidungsprozesse nicht unbedingt ungenauer als normative Entscheidungsprozesse oder ihnen unterlegen sind.

Wie unser Gehirn Entscheidungen trifft

Die Neuroökonomie ist eine Disziplin, die Neurowissenschaften und Wirtschaftswissenschaften vereint und mittels neurowissenschaftlicher Methoden unter anderem Gehirnaktivitäten in bestimmten Entscheidungssituationen untersucht (Sanfey 2007). Die Ergebnisse dieser noch jungen

und spannenden Wissenschaft liefern erhellende Aufschlüsse über menschliche Entscheidungsprozesse und stützen etliche Entscheidungstheorien in überraschender Weise.

Wie wir gesehen haben, beschreibt die Theorie der rationalen Entscheidung – die Urtheorie aller Entscheidungstheorien – einen rationalen Entscheidungsprozess als das Produkt aus Wahrscheinlichkeitsschätzung und Nutzen. Es stellt sich heraus, dass diese beiden Funktionen neurologisch real und neurologisch separierbar sind. Der Teil Ihres Gehirns, der die Wahrscheinlichkeit einer Entscheidung berechnet, operiert separat von dem Teil, der berechnet, wie glücklich Sie diese Entscheidung machen wird.

Beginnen wir mit dem Nutzen. Dopamin ist ein Neurotransmitter (ein biochemischer Botenstoff), der freigesetzt wird, wenn man eine Belohnung erhält oder sich in einer Phase der Antizipation, der Vorfreude, auf eine Belohnung, befindet (Box 3.2). Selbst Kriterien, die mit einer Belohnung verbunden sind, lösen eine Freisetzung von Dopamin aus. Neuronen, die Dopamin ausschütten, kommen in vier Hirnarealen vor: dem Nucleus accumbens, der Area tegmentalis ventralis (*ventral tegmental area*, VTA), dem Striatum und dem frontalen Cortex. Diese Areale gelten als Belohnungssystem unseres Gehirns. Alles, was uns Spaß und Freude bereitet (ob wir unsere Lieblingsmusik hören oder ein schönes Gesicht sehen), aktiviert dieses System. Diese Schaltkreise versetzen unser Gehirn in die Lage, die Umstände, die uns Freude bereitet haben, zu codieren und sich an sie zu erinnern, damit wir dieses Verhalten in der Zukunft wiederholen und die Belohnung wieder abrufen können. Sind diese Schaltkreise aktiviert, so sind, wie man auch sagen könnte, die neuronalen Signaturen für die Belohnungsverarbeitung (und damit für die Verarbeitung des

Nutzens) in vollem Gange. Egal, ob Sie eine Belohnung erhalten oder eine Entscheidung treffen, von der Sie glauben, Sie bringe Ihnen eine Belohnung, es werden immer dieselben Areale aktiviert. Insofern macht Ihr Gehirn also keinen Unterschied zwischen erfahrenem Nutzen (*experienced utility*) und Entscheidungsnutzen (*decision utility*) (Breiter et al. 2001; O'Doherty et al. 2002).

Box 3.2 Angriff auf das Belohnungssystem im Gehirn

Alle Suchtmittel stimulieren die Dopaminausschüttung im Belohnungssystem des Gehirns sehr viel mehr als es im alltäglichen Leben normalerweise vorkommt. Bei fortgesetztem Suchtmittelmissbrauch kommt es zu einer wahren Dopaminflut, auf die das Gehirn reagiert, indem es weniger Dopamin produziert oder die Zahl der Dopaminrezeptoren im Belohnungssystem reduziert. Dies wiederum führt dazu, dass die weitergeleiteten Signale, die dem Drogenkonsumenten Glücksgefühle bescheren, schwächer sind als sonst. Der Konsument empfindet ein Belohnungsdefizit und verspürt nun das Verlangen, die Dosis zu erhöhen, um das Dopaminsystem wieder in sein altes Gleichgewicht zu bringen und sich wieder „high" zu fühlen.

Beispiel Kokain: Das konsumierte Kokain blockiert, nachdem ein Neuron erregt wurde, die Wiederaufnahme von Dopamin aus dem synaptischen Spalt in das präsynaptische Neuron. Kokain ist also ein Wiederaufnahmehemmer und sorgt damit dafür, dass das ausgeschüttete Dopamin länger im synaptischen Spalt verweilt, wodurch die Rezeptoren länger stimuliert und Glücksgefühle ausgelöst werden. Diese Überstimulierung führt mit der Zeit zu einer Schädigung oder Zerstörung der Dopaminrezeptoren, und ihre Zahl verringert sich. Die Folge: Der Süchtige braucht immer größere Drogenmengen, um den gewünschten Kick zu bekommen und das Belohnungssystem aktiv zu halten.

Für Wilson und Kuhn (2005) ist Sucht weit mehr als nur die (verhängnisvolle) Suche nach einem Zustand der Eupho-

rie, verbunden mit dem großen Verlangen, Entzugserscheinungen möglichst zu vermeiden. Der exzessive Gebrauch von Drogen ist ein Angriff auf die neuronalen Schaltkreise im Gehirn und treibt den Süchtigen direkt in ein Abhängigkeitsverhalten. Bei fortgesetztem Konsum unterwirft sich das Belohnungssystem des Gehirns dem Suchtverlangen.

Mithilfe der funktionellen Magnetresonanztomografie (fMRT oder auch fMRI), einem bildgebenden Verfahren, das aktivierte Hirnareale darstellen kann, führten Knutson et al. (2005) ein Experiment durch, in dem sowohl der Größenwert der Belohnung als auch die Wahrscheinlichkeit für das Auftreten eines Ereignisses manipuliert wurden. Die Teilnehmer bekamen Hinweise auf beides präsentiert, also einen Wahrscheinlichkeitswert ebenso wie einen Größenwert der bevorstehenden geldlichen Belohnung. Die Forscher fanden heraus, dass Aktivitäten im medialen präfrontalen Cortex geknüpft waren an die subjektive *Wahrscheinlichkeit*, die Belohnung zu erhalten, während Aktivitäten in Arealen des Mittelhirns mit dem zu erwartenden *Größenwert* der Belohnung korrelierten. Darüber hinaus zeigte sich, dass die mündlichen Berichte der Teilnehmer über ihre jeweils subjektiven Einschätzungen der Wahrscheinlichkeiten mit Aktivitäten in präfrontalen Gehirnarealen korrelierten, die Berichte über ihren jeweiligen Erregungszustand hingegen mit Aktivitäten im Mittelhirn. Weitere Studien, die ebenfalls unter Verwendung bildgebender Verfahren durchgeführt wurden (fMRI: Breiter et al. 2001; EEG: Holroyd et al. 2004), zeigten ebenfalls, dass sich die neuronalen Signaturen für die Belohnungsverarbeitung nach dem zu erwartenden Ergebniswert richten, und zwar relativ zur Bandbreite aller möglichen Ergebnisse.

Sie richten sich also nicht nach dem objektiven Wert des Ergebnisses selbst, wie es nach der Neuen Erwartungstheorie (*prospect theory*) der Fall sein müsste.

Ergebnisse wie diese zeigen sehr schlüssig, dass Wahrscheinlichkeitseinschätzungen auf der neurologischen Ebene separierbar sind von der (subjektiven) Beurteilung der Wertigkeit der Belohnung. Die Studie zeigte außerdem, dass die Modelle der Nutzentheorie (*utility theory*) ein sehr genaues Bild darüber vermitteln können, wie das Gehirn zwischen zwei möglichen Alternativen entscheidet. Der entscheidende Unterschied jedoch ist der, dass ökonomische Standardmodelle, die auf der Theorie der rationalen Entscheidung basieren, einen einzigen Prozessor für die rationale Informationsverarbeitung voraussetzen. Neurowissenschaftliche Forschungsergebnisse deuten aber daraufhin, dass Entscheidungen das Resultat zweier separat ablaufender neuronaler Prozessoren sind.

Gleichwohl fanden die Forscher heraus, dass diese beiden separaten Prozessoren in Entscheidungssituationen miteinander konkurrieren. Wenn wir beschließen, eine risikobehaftete Entscheidung zu treffen, so sind unsere Belohnungszentren im Gehirn unmittelbar bevor die Entscheidung erfolgt hoch aktiv. Mit anderen Worten, diese neuronale Signatur zeigt, dass wir hohe Auszahlungen antizipieren und nicht über die *Wahrscheinlichkeit* der Auszahlungen nachdenken (Knutson & Bossaerts 2007). Dies ist ein Grund, warum Glücksspiele süchtig machen können; allein eine Wette zu platzieren, regt das Belohnungssystem an und kann ein ebenso großes Glücksempfinden auslösen wie ein Gewinn selbst.

Die neuronalen Signaturen, die Framing-Effekten zugrunde liegen, wurden ebenfalls identifiziert. Um den Me-

chanismus der Framing-Effekte eingehender zu untersuchen, ließen De Martino et al. (2006) die Probanden einer Studie Wetten eingehen, die eine unterschiedliche „Rahmung" (*frame*) hatten, sprich, die unterschiedlich formuliert waren, und zwar als sichere oder unsichere Gewinne oder Verluste (wie oben erörtert). Sie fanden heraus, dass emotionsverarbeitende Areale des Gehirns (wie der Mandelkern, auch Amygdala genannt) bei den Teilnehmern, die solchen mentalen Rahmungseffekten anheimfallen, überaus aktiviert sind. Bei Teilnehmern, die diesen Rahmungseffekten nicht erliegen, zeigen diese Areale eine verminderte Aktivität, der präfrontale Cortex (insbesondere der anteriore präfrontale Bereich, der einen Konflikt zwischen den Ergebnissen des emotionsbasierten und denen des verstandesbasierten Systems signalisiert) hingegen zeigt eine gesteigerte Aktivität. Diese Ergebnisse, so hoben die Forscher hervor, deuteten auf eine Konkurrenz zwischen zwei neuronalen Systemen, die an der Aktivierung des anterioren cingulären Cortex (ACC), des emotionsbasierten Systems (Amygdala) sowie des analytischen Systems (orbitofrontaler und präfrontaler Cortex) zu erkennen ist. Des Weiteren zeigte sich, dass die Teilnehmer, bei denen eine gesteigerte Aktivierung im orbitofrontalen und präfrontalen Cortex zu erkennen war, auch am wenigsten von der Manipulation durch Framing-Effekte beeinflusst waren.

Und noch eine Studie: 2008 legten De Neys und Kollegen ihren Probanden Kurzbeschreibungen von Prototypen eines Ingenieurs oder eines Juristen vor, oder die neutral gehalten und damit nicht konkurrenzbasiert waren. Die Probanden sollten nun entscheiden, mit welcher Wahrscheinlichkeit es sich um einen Ingenieur bzw. Juristen handelte. Die Ergebnisse zeigten, dass der ACC in der kon-

kurrenzbasierten Entscheidungssituation in der Tat aktiver war als in der neutralen Entscheidungssituation. Sofern die Probanden es schafften, ihren ersten Impuls zu überwinden und sich in ihren Entscheidungen eben nicht auf den Prototypen festzulegen, mit dem die größte Ähnlichkeit vorlag, war der laterale präfrontale Cortex überaus aktiv. Und das bedeutet, dass sich vorschnelle Entscheidungsfindungen scheinbar dann ergeben, wenn es nicht gelingt, intuitive Heuristiken außer Kraft zu setzen. In besagter Studie trifft System 1 also eine rasche Entscheidung, indem es eine Beschreibung blitzschnell mit den angelegten Prototypen für Ingenieure und Juristen abgleicht. Was immer am besten passt, ist die richtige Antwort. System 2 gelangt sehr viel langsamer zu einer Entscheidung, indem es sich auf die Struktur der Aufgabe fokussiert, insbesondere auf die gegebenen Wahrscheinlichkeitsinformationen. System 2 gründet seine Entscheidung also auf vorliegenden Kenntnissen der Wahrscheinlichkeiten. Waren die letzten Entscheidungen der Probanden heuristisch basiert (ähnlichkeitsbasiert), gewann System 1 den Entscheidungswettbewerb. Waren sie hingegen vernunftbasiert, hatte System 2 die Nase vorn.

Entscheidungssituationen in der Realität: die Weltwirtschaftskrise 2008

Man muss keine gedanklichen Klimmzüge anstellen, um festzustellen, dass überbordende Eigeninteressen zu den ursächlichen Faktoren der globalen Finanzkrise gehörten, die

2008 die ganze Welt in die Zange nahm. Wie konnte es dazu kommen? Laut dem offiziellen Untersuchungsbericht der US-Untersuchungskommission zur weltweiten Finanzkrise (Financial Crisis Inquiry Commission) von 2011 war Folgendes passiert: Banken gaben Hypothekenanträgen ihrer Kunden leichthin statt, da auf dem Geldmarkt viel und billiges Geld vorhanden war und sie so profitable Geschäfte witterten. Sie bündelten diese Hypotheken, verkauften sie als sichere Kapitalanlagen weiter und entledigten sich so jeglicher Risiken oder Verantwortlichkeiten, sollten die Hypothekennehmer (Schuldner) die Kredite nicht zurückzahlen können. Immobilienmakler auf den Finanzmärkten machten mit jedem Haus, das sie verkauften, satte Gewinne, weshalb es in ihrem ureigennützigen Interesse lag, so viele Häuser wie möglich zu verkaufen, ob die Käufer am Ende zahlungsfähig waren oder nicht. Mit Hypotheken hatten sie schließlich nichts zu schaffen, das war Problem der Bank. Da Immobilienkredite günstig zu haben waren, nahmen die Menschen höhere Kredite auf, als sie sich leisten konnten, in der Hoffnung, ihr Haus gewinnbringend wieder verkaufen oder es späterhin zu einer niedrigeren Rate refinanzieren zu können. Inzwischen stiegen die Preise für Immobilien überproportional an, denn das Kaufinteresse war riesig und löste regelrechte Bieterkriege aus.

Das funktionierte so lange, bis die Blase platzte. Die Menschen, die eine Hypothek erhalten hatten, sie sich finanziell aber gar nicht leisten konnten, waren nun zahlungsunfähig. Das traf insbesondere die, deren Hypothekenverträge mit variablen Zinssätzen versehen waren, die am Anfang für eine bestimmte Zeit attraktiv niedrig ausfielen und später an das Niveau eines durchschnittlichen Marktzinses ge-

koppelt waren. Wer anfangs beispielsweise eine monatliche Rate von 500 Dollar auf das Heim zu zahlen hatte, musste nun plötzlich 1500 Dollar monatlich berappen. Und das konnten sich viele nicht mehre leisten. Also inserierten sie es zum Verkauf, wie es Millionen anderer Menschen machten, die sich ihre Immobilien ebenfalls nicht mehr leisten konnten.

Immer mehr Immobilien überschwemmten den Markt, während gleichzeitig immer mehr Menschen nach Finanzierungs- oder Refinanzierungsmöglichkeiten suchten: Die Immobilienblase platzte. Die Banken begannen, Häuser wegen nicht bezahlter Hypotheken zu pfänden und an Menschen, die keine Kredite erhielten, da sie keine Sicherheiten zu bieten hatten, zu versteigern. Dies setzte eine Kettenreaktion in der Realwirtschaft in Gang.

Viele Banken und Investmentfirmen waren infolge massiver Verluste aus hypothekenbesicherten Geschäften schwer angeschlagen. Der Bankenriese Lehman Brothers ging am 15. September 2008 Pleite und andere schienen unaufhaltsam zu folgen. Diese hypothekenbesicherten Sicherheiten, denen Privatimmobilien als Sicherheit dienten, waren weltweit vermarktet worden, weshalb sich die Finanzkrise nicht nur auf die USA beschränkte. Dies führte letztlich zu einer der schlimmsten Rezessionen, wie es sie seit der Großen Depression in den 1930er-Jahren nicht mehr gegeben hatte – und der Zusammenbruch der weltweiten Finanzmärkte, angestiftet von falschen Anlageentscheidungen, hatte begonnen. Ein Werteverfall US-gesicherter Kapitalanlagen bereitete der amerikanischen Regierung zunehmend Sorge, die diese Finanzinstitute, welche mit ihrem Scheitern nun die ganze Welt in den Abgrund zu reißen drohten, für

too big to fail („zu groß, um zu scheitern") erachtet hatte. Um das System zu stabilisieren, organisierte die US-Regierung die größte staatliche Bankenrettung der Geschichte. Die Steuergelder einfacher Bürger wurden verwendet, um das sinkende Schiff der US-Investmentbankindustrie über Wasser zu halten.

Gewiss, Eigennutz und Profitgier mögen die Wurzeln dieser Wirtschaftskrise begründet haben, was allerdings der Tatsache geschuldet ist, dass die Akteure, die unsere Märkte bestimmen, Menschen sind, und wie wir gesehen haben, verträgt sich die Psychologie der menschlichen Entscheidungsfindung mit den Kräften des Marktes katastrophal schlecht. Wir müssen heute häufig Entscheidungen treffen, die auf Kalkulationen basieren, welche wir im Kopf nur schwer durchführen können. Hinzu kommt ein Unsicherheitsmoment, das heißt wir fällen viele Investitionsentscheidungen in einer Unsicherheit, da wir mögliche Ergebnisse oder Entwicklungen unserer Entscheidungen nicht mit Sicherheit bestimmen können. Und das wiederum liegt daran, dass wir entweder nur unzureichende Informationen haben oder die Informationen asymmetrisch verteilt sind, will heißen, die andere interessierte Partei weiß mehr über eine Investition als wir selbst. Wie P. J. O'Rourke in seinem Buch *Eat the rich* herausstellt, ist es sehr schwierig, asymmetrische Informationen und Insider-Handel auseinanderzuhalten, da es genau dasselbe ist. Auf diese Weise landete Martha Stewart am Ende im Gefängnis. (Die prominente amerikanische Fernsehmoderatorin und Unternehmerin wurde wegen illegaler Aktiengeschäfte angeklagt, da sie in den Verdacht geriet, auf Grund von Insider-Informationen gehandelt zu haben; Anm. d. Übers.) Wir versuchen, Ri-

siken zu vermeiden, doch treffen wir hochriskante Entscheidungen, wenn wir uns von potenziellen Verlusten bedroht sehen. Wenn wir eine Kapitalanlage erwerben, die daraufhin prompt im Wert steigt, sind wir häufig geneigt, sie wieder abzustoßen, um den Gewinn einzustreichen. Wenn wir hingegen eine Kapitalanlage erwerben, die daraufhin prompt im Wert sinkt, sind wir häufig geneigt, sie auf unbestimmte Zeit zu halten in der Hoffnung, sie werde sich eines Tages erholen und zurück auf ihren ursprünglichen Wert klettern. Wir zocken sozusagen bis zum bitteren Ende. Zudem neigen wir nicht selten zu einem übersteigerten Selbstvertrauen, das der Besonnenheit in komplexen Entscheidungssituationen oft ein Schnippchen schlägt.

In Anbetracht all dessen gibt es heute in der finanzwirtschaftlichen Fachwelt etliche Strömungen, die in ihren Beschreibungen der Märkte auch die Strukturen der menschlichen Psyche in den Blick nehmen. Dieses Teilgebiet der Wirtschaftswissenschaften wird als Verhaltensökonomie bezeichnet. Ihre Forschungen machen Hoffnung, dass sich nicht nur wirtschaftliche Modellannahmen zum menschlichen Entscheidungsverhalten verbessern lassen, sondern auch die eigenen Fähigkeiten, die wir brauchen, um potenziell negative Konsequenzen aus konkreten Entscheidungen zu verhindern.

4

Moralische Urteilsbildung

Wie wir Richtig von Falsch unterscheiden

1978 brachte die Philosophin Philippa Foot erstmals das folgende Gedankenexperiment zu einem moralischen Dilemma vor, das unter dem Namen „Trolley-Problem" bekannt ist:

> Eine Straßenbahn ist außer Kontrolle geraten und droht, fünf Personen zu überrollen, die von einem geisteskranken Philosophen an die Gleise gekettet wurden. Zum Glück können Sie eine Weiche umstellen, um die Straßenbahn auf ein anderes Gleis umzuleiten und so das Unglück abzuwenden. Unglücklicherweise befindet sich auch dort eine weitere Person, die an die Gleise gekettet ist. Was tun Sie? Würden Sie die Weiche umstellen oder würden Sie nichts tun?

Die meisten der Befragten würden die Weiche umstellen und damit einen Menschen opfern, um fünf andere zu retten. Das erscheint ihnen richtig. Doch betrachten wir nun die folgende Variante des Trolley-Problems, die Judith Jarvis Thomson (1985) vorbringt:

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_4,

> Eine außer Kontrolle geratene Straßenbahn rollt auf fünf Personen zu. Sie selbst stehen auf einer Brücke, unter der die Straßenbahn durchfährt, und können sie aufhalten, indem Sie ihr einen schweren Gegenstand in den Weg werfen. Nur leider ist weit und breit nichts zu sehen außer einem sehr dicken Mann, der direkt neben Ihnen steht – die einzige Möglichkeit, die Straßenbahn zu stoppen, ist, ihn über die Brücke auf die Gleise zu stoßen, ihn also zu töten, um fünf andere Menschenleben zu retten. Würden Sie es tun?

Im Gegensatz zur Standardversion, wie sie Philippa Foot formuliert, glauben die meisten Befragten nun, dass es falsch sei, den dicken Mann von der Brücke zu stoßen – auch wenn dadurch fünf Menschenleben geopfert würden.

Wie diese Problemstellungen zeigen, scheint unsere Intuition bisweilen sehr widersprüchlich, wenn es um moralische Fragen geht, die obendrein sehr starke emotionale Reaktionen in uns auslösen und weitreichende Auswirkungen auf das Leben anderer Menschen haben können. Der amerikanische Bürgerkrieg beispielsweise wurde zu einem großen Teil auch deshalb ausgetragen, weil sich das Land in der moralischen Frage der Sklaverei uneins war. Und erst als die Menschen zu der Überzeugung gelangten, dass sozialpolitische Ungleichheiten wie sie in den sogenannten Jim-Crow-Gesetzen, den Gesetzen zur Rassentrennung, festgeschrieben waren, moralisch falsch sind, konnte die US-amerikanische Bürgerrechtsbewegung (*civil rights movement*) unaufhaltsam an Dynamik und Bedeutung gewinnen. Im Wissen um den großen Einfluss, den moralische Dilemmas mit ihrer oft schwierigen und komplexen Natur auf unser Leben haben, suchen wir in diesen Fragen häu-

fig nach Orientierungshilfen. Dieses Kapitel erhebt keinen Anspruch, eine umfassende Abhandlung dieses so gewichtigen Themas zu liefern. Es soll vielmehr darum gehen, die einflussreichsten Betrachtungen von Theoretikern der säkularen und ethischen Philosophie der westlichen Kultur in kurzen Abrissen vorzustellen.

Kirche, Staat und Moral

Zunächst sei bemerkt, dass die Theorie der rationalen Entscheidung hier nicht wirklich brauchbar ist. Wie wir gesehen haben, liegt der Grundsatz dieser Theorie darin, dass rational handelnde Subjekte (Akteure) stets in ihrem eigenen Interesse agieren. Eigeninteressen aber scheinen hier nicht auf einer Linie mit dem Trolley-Problem zu liegen. Vielmehr schließt das Dilemma fremde Personen ein, das eigene Leben ist nicht bedroht, und von daher ziehen wir selbst offenbar keinen unmittelbaren Nutzen aus der Entscheidung, egal welche wir treffen.

Probleme dieser Art bauen stattdessen auf das implizite Prinzip eines *moralischen Imperativs*, auf eine Handlung, die man auszuführen hat, da genau so zu handeln richtig ist. Während ich diese Zeilen niederschreibe, habe ich bei Google das Stichwort „moralischer Imperativ" eingegeben und über eine Million Treffer erhalten, meist eingebunden in kleinere Überschriften oder Halbsätze:

Moralischer Imperativ und Obama
Moralischer Imperativ und schulische Bildung
Moralischer Imperativ und schockierende Forschungsergebnisse

Moralischer Imperativ und das Töten libyscher Islamisten
Moralischer Imperativ und Bekämpfung der Armut

Ganz gleich, auf welches Thema sich der moralische Imperativ beziehen mag, er ist nicht bloß als Vorschlag zu verstehen, den zu befolgen wünschenswert oder erstrebenswert wäre – nein. Er ist in seiner Formulierung verpflichtend. Es ist folglich nicht nur eine gute Entscheidung, die schulische Bildung auszubauen und zu verbessern, es ist eine moralische Pflicht. Dieser Pflicht nicht nachzukommen, wird als moralisch falsch beurteilt.

Moralische Dilemmas wie das Trolley-Dilemma weisen jedoch eine weitere Besonderheit auf: In allen Handlungsalternativen, die zur Verfügung stehen, konkurriert ein moralischer Imperativ mit einem anderen; den einen zu befolgen, würde heißen, gegen den anderen zu verstoßen. Es ist also egal, für welche Alternative man sich entscheidet, ein Moralprinzip wird man immer brechen. Ob in der Standardversion des Trolley-Problems oder in der Variante mit dem dicken Mann – in beiden Fällen gilt es zu entscheiden, ob man ein Menschenleben opfern will, um viele weitere zu retten, oder nicht. Und genau an diesem Punkt geraten zweierlei moralische Imperative aneinander. Der eine besagt, dass es moralisch richtig sei, Menschenleben zu retten. Der andere besagt, dass es moralisch falsch sei, Menschen zu töten. Doch unsere widerstreitenden Reaktionen auf derlei Dilemmas legen nahe, dass sich darin mehr moralische Imperative versteckt halten, als sich auf den ersten Blick erkennen lassen. Aber welche moralischen Imperative sind das, und wie lassen sie sich erklären?

Die Antworten auf diese Fragen hängen weitgehend davon ab, ob man eine religiöse oder eine säkulare Sichtweise einnimmt. Religionen schreiben moralische Imperative üblicherweise einer göttlichen Autorität zu. Die *Theokratie* ist eine Herrschaftsform, in der die Ausübenden von Staatsgewalten als gottberufen angesehen werden, weshalb die durch sie verhängten Gesetze moralische Kraft haben. Eine Theokratie wird bisweilen auch als „(ge)offenbarte Religion" bezeichnet – das heißt, durch Gott werden dem Menschen Ideen offenbar, die durch den menschlichen Verstand nie erlangt werden können. Der Bildung einer Regierung auf einem solchen Fundament standen die Gründungsväter der Vereinigten Staaten von Amerika überaus skeptisch gegenüber. Thomas Jefferson schrieb 1802 in einem Brief an eine Gruppe, die sich selbst als Danbury Baptists bezeichnete, folgende Zeilen:

> „Mit Euch halte ich es für wahr, dass Religion eine Angelegenheit ist, die allein den Menschen und seinen Gott betrifft, dass niemand einem anderen Rechenschaft schuldig ist für seinen Glauben und seine Anbetung und dass sich die legitimierte Gewalt der Regierung allein auf Handlungen erstreckt und nicht auf Meinungen. So betrachte ich mit souveräner Ehrfurcht das Gesetz des amerikanischen Volkes, das verkündet, dass seine Legislative ‚kein Gesetz erlassen darf, das die Einrichtung einer Religion betrifft oder ihre freie Ausübung verbietet', und hiermit eine Trennungsmauer zwischen Kirche und Staat einrichtet."[1]

[1] Van der Ven J (2010) Trennung von Kirche und Staat. 63–112; S. 80 In: Ziebertz H-G (Hrsg) Menschenrechte, Christentum und Islam. Lit, Münster

Jefferson war ebenso wie John Adams und James Madison sehr stark beeinflusst von den Philosophen der Aufklärung des 18. Jahrhunderts, die Vernunft und wissenschaftliche Beobachtung als wichtigste Werkzeuge betonten, um die Wahrheit der Dinge zu erkennen. Der einflussreichste unter ihnen war David Hume. Mit seinen Ideen setzten sich insbesondere drei maßgebende Philosophen des 19. Jahrhunderts kritisch auseinander – Immanuel Kant, John Stuart Mill und Jeremy Bentham. Die Werke dieser drei herausragenden Denker bilden zusammengenommen das grundlegende Fundament der modernen Moralphilosophie oder Morallehre (Ethik).

Die Moralphilosophie beschäftigt sich mit Fragestellungen, die sich um Grundbegriffe wie „gut" und „böse", „richtig" und „falsch", „Tugend" und „Laster" oder „Recht" und „Gerechtigkeit" drehen. Dabei geht es nicht darum, wie wir uns *tatsächlich* verhalten, sondern vielmehr darum, wie wir uns verhalten *sollen*. Im Kern sind damit drei moralphilosophische Prinzipien verbunden: Moralisch erlaubte Handlungen sind Handlungen, die wir ausführen dürfen, wenn wir es wollen (z. B. für wohltätige Zwecke spenden). Moralisch verbotene Handlungen sind Handlungen, die verboten und damit falsch sind (z. B. einen unschuldigen Menschen töten). Moralisch gebotene Handlungen sind Handlungen, die wir ausführen *sollen*, sofern wir dazu in der Lage sind, und die zu unterlassen falsch ist (z. B. das Leben einer unschuldigen Person retten). Diese Prinzipien sind auch die Eckpfeiler der Rechtstheorie.

Was Hume zu sagen hatte

David Hume (1711–1776) war ein schottischer Philosoph, Ökonom und Historiker. Er war Gründer einer philosophischen Strömung, die dem britischen Empirismus zugerechnet wird und die den Absolutheitsanspruch wissenschaftlicher Erkenntnis ablehnt. Der Empirismus geht davon aus, dass alles Wissen allein durch (Sinnes-)Erfahrungen erlangt werden kann. Hume legt seine sehr einflussreiche Moraltheorie in Band 3 seines Werks *Ein Traktat über die menschliche Natur* (1740) sowie in *Untersuchung in Betreff des menschlichen Verstandes* (1751) dar. Im ersten Teil von Band 3 stellt er folgende Frage (die bis heute eine Kernfrage wissenschaftlicher Untersuchungen und philosophischer Abhandlungen ist): Sind moralische Urteile auf Vernunft gegründete Urteile über konzeptionelle Beziehungen und einzelne Tatsachen oder sind sie emotionale Reaktionen? Hume war überzeugt, es seien emotionale Reaktionen. Er bringt sein berühmtes „Argument vom Vatermord" vor, um seine Position zu stützen: Ein junger Baum, der seine Baum-Eltern überwuchert und damit tötet, weist die gleiche behauptete Beziehung auf wie ein menschliches Kind, das seine Eltern tötet. Würden Moral und moralische Urteile also lediglich auf erkennbaren Beziehungen basieren, so wäre der junge Baum unmoralisch. Und das, so stellt Hume heraus, sei schlichtweg absurd. Der wesentliche Punkt in Humes Analyse ist folgender: Wir können aus Aussagen über Tatsachen (Istaussagen) keine Aussagen über moralische Verpflichtungen (Sollaussagen) ableiten. Und da moralische Gutheißungen nicht auf Vernunft gegründet sind, müssen sie folglich eine emotionale Reaktionen sein.

Hume entwickelt eine Moraltheorie, in der die Handlungskette von den Beteiligten selbst bestimmt wird – vom Handelnden selbst, vom Empfänger dieser Handlung und von einem außen stehenden Beobachter. Hume geht davon aus, dass moralische Handlungen ihren Wert nur durch charakterliche Motive des Handelnden erhalten, die entweder tugendhaft oder boshaft sein können. Zum Beispiel sind Menschen, die freiwillig Geld für wohltätige Zwecke spenden, durch eine tugendhafte Eigenschaft motiviert. Hingegen sind diejenigen, die Geld stehlen, um sich selbst zu bereichern, durch eine boshafte Eigenschaft motiviert.

Hume glaubt, dass einige Charaktereigenschaften in der Natur des Menschen festgelegt sind, andere wiederum erworben werden, also „künstliche Tugenden" sind, um seine Worte zu gebrauchen. Um noch einmal auf unser Spendenbeispiel zu kommen: Für wohltätige Zwecke zu spenden, gibt uns ein gutes Gefühl. Diese Art der moralischen Handlung ist also Ausdruck eines natürlichen Gefühls. Der Empfänger dieser Handlung mag Dankbarkeit empfinden, was ebenfalls Ausdruck eines positiven Gefühls ist. Und schließlich kann der Beobachter das positive Gefühl des Empfängers nachempfinden, wenn er die Handlung der Nächstenliebe beobachtet. Würde man Spendengelder jedoch stehlen, würde man sich schlecht fühlen, genau wie die Bestohlenen selbst und jeder, der den Diebstahl beobachtet. *Hume postuliert, dass die Empfindungen der Beteiligten den moralischen Stellenwert einer Handlung bestimmen.*

Er bringt in seine Moraltheorie durchaus auch eine utilitaristische Idee ein: Wir heißen moralische Handlungen teilweise deshalb gut, weil sie ihre Grundlage in einem *Nutzen* haben – sie also *nützlich* sind. In seinem Werk *Untersuchung über die Prinzipien der Moral* (1751) argumentiert

Hume in Abschnitt V unter der Überschrift „Warum gefällt die Nützlichkeit?", dass wir solch nützliche Handlungen deshalb gutheißen, da sie „aus einer so edlen Quelle wie dem Mitgefühl herrühren".[2]

Indem er moralische Urteile in Gefühlen begründet sieht, erklärt seine Theorie mithin auch, warum unsere Reaktion auf die Standardversion des Trolley-Problems anders ausfällt als auf die Variante mit dem dicken Mann: Die hautnahe und persönliche Note in der Variante mit dem dicken Mann löst stärkere Gefühle aus als die Standardversion mit der ferngesteuerten Weichenumstellung. Einen Schalter umzulegen und damit einen zwar unglücklichen, aber dennoch günstigen Ausgang zu erwirken, ist eine Sache, eine ganz andere ist es, einen Menschen gegen seinen Willen direkt zu packen und ihn in den Tod zu stoßen. Auch wenn in beiden Versionen dieselbe Anzahl Menschen geopfert bzw. gerettet wird, könnte die emotionale Reaktion auf ein und dasselbe Problem unterschiedlicher nicht sein.

Hume beschäftigte sich überdies auch mit dem etwas abstrakteren und politischen Begriff der Gerechtigkeit. Nach seinem Dafürhalten gründet die Gerechtigkeit (der Rechtssinn) nicht in der Natur des Menschen, sondern sie entsteht durch gesellschaftliche Vereinbarungen und Konventionen und wird durch Erziehung weitergegeben. Da wir auf die Gesellschaft angewiesen sind, um zu überleben, entspricht es dem menschlichen Interesse, sie aufrechtzuerhalten, indem wir eine Rechtsordnung etablieren und diese sichern. Es gilt daher, unsere Verantwortung für andere in der Gesellschaft anzuerkennen, um dieses Ziel zu erreichen.

[2] http://www.neuemoral.de/www_neuemoral_de/Philosophen/David_Hume/david_hume.html (Zugriff 8.1.2014)

Zu den drei wichtigsten Rechtsregeln, die sich aus diesen Betrachtungen ergeben, gehören: Anerkennung von Eigentum (als Grundlage für eine stabile Gesellschaft), Übertragung von Eigentum in gegenseitigem Einvernehmen, Leistungsversprechungen (Verträge). Regierungen bilden sich heraus, um die Anwendung oder Erfüllung eingegangener Vereinbarungen zu sichern und uns so als Mitglieder der Gesellschaft sowie die Gesellschaft als Ganzes zu schützen.

Nach Hume ist das höchste Verdienst, das die menschliche Natur zu erreichen fähig ist, wohlwollende Menschlichkeit. Um dieses Kapitel abzuschließen, will ich hier Abschnitt 2 „Von dem Wohlwollen" aus *Ein Traktat über die menschliche Natur* zitieren. Die Ideen, die hier geäußert werden, geben Ihnen einen feinen Eindruck dieses machtvollen Aufklärungsgedankens über Gerechtigkeit und die menschliche Natur. Beachten Sie, dass Hume nicht nur an Anstand und Fairness appelliert, sondern an *noblesse oblige*:

> „Der Beweis, daß unsere wohlwollenden oder sanftmütigeren Gefühle schätzenswert sind und daß sie immer die Billigung und den guten Willen der Menschheit hervorrufen, darf vielleicht als ein überflüssiges Unternehmen angesehen werden. Die Attribute *gesellig*, *gutmütig*, *menschlich*, *gütig*, *dankbar*, *freundlich*, *großzügig*, *wohltätig* oder deren Entsprechungen sind in allen Sprachen bekannt, und sie drücken universell das höchste Verdienst aus, das die *menschliche Natur* zu erreichen fähig ist. Wenn diese liebenswerte Eigenschaften mit hoher Geburt, Macht und herausragenden Fähigkeiten verbunden sind und sich in guter Regierung oder nützlicher Belehrung der Menschheit zeigen, dann scheinen sie sogar ihre Eigentümer über den Rang der *menschlichen Natur* hinauszuheben und sie

in gewisser Hinsicht der göttlichen anzunähern. Höchste Begabung, unbezwingbarer Mut, großer Erfolg können einem Helden oder einem Politiker vielleicht nur den Neid und die Feindschaft der Öffentlichkeit einbringen, aber sobald das Lob von Menschlichkeit und Wohltätigkeit hinzukommen, wenn Beispiele von Nachsichtigkeit, Milde oder Freundschaft aufgezeigt werden, dann wird der Neid still oder schließt sich dem allgemeinen Ausdruck von Billigung und Beifall an.
Als Perikles, der große athenische Staatsmann und General auf seinem Totenbett lag, begannen die Freunde, die um ihn waren und ihn als nicht mehr aufnahmefähig betrachteten, ihre Trauer über ihren sterbenden Patron auszudrücken, indem sie seine großen Eigenschaften und Erfolge, seine Eroberungen und Siege, die ungewöhnliche Länge seiner Regierungszeit und die neun Trophäen, sie als Zeichen seines Triumphs über die Feinde der Republik errichtet worden waren, aufzählten. Der sterbende Held aber, der alles gehört hatte, rief: ‚Ihr vergeßt das wichtigste Lob, indem ihr so sehr auf diese vulgären Vorzüge Acht habt, in denen das Glück einen wichtigen Anteil hatte. Ihr habt nicht bemerkt, daß bis jetzt kein Bürger meinetwegen jemals ein Trauerkleid trug.“[3]

Was Kant zu sagen hat

Der deutsche Philosoph Immanuel Kant widersprach der Sichtweise der Empiriker, wonach alles Wissen aus der Erfahrung resultiert. Stattdessen argumentiert er, dass die Vernunft der Quell aller Erkenntnisse und Begründbarkeiten

[3] Zitat aus: Plutarch, Perikles (Perikles *Leben* 38 §4, 173C)

sei. Und dies gelte, so Kant, auch für die Moral. In scharfem Gegensatz zu Humes Überzeugung, dass Gefühle die Grundlage für moralisches Handeln seien, postuliert Kant eine Moraltheorie, der zufolge *moralische Urteile das Ergebnis rationaler Gedanken sind*. Kants Moraltheorie wird als Deontologie (deontologische Ethik) bezeichnet, als Pflichtenlehre, bei der die Pflichten (Rechte und Verpflichtungen) des moralisch Handelnden im Vordergrund stehen. Seine Theorie der Moral entwickelt Kant in seinen Werken *Grundlegung zur Metaphysik der Sitten* (1785) und *Kritik der reinen Vernunft* (1787). Grundlegendes Prinzip seiner Theorie ist der *kategorische Imperativ* – er gebietet Handlungsregeln (Maximen), die sich allein aus der Vernunft herleiten lassen und für alle vernünftig handelnde Wesen absolut verpflichtend sind.

Wie Hume glaubt auch Kant, dass die moralische Bewertung einer Handlung abhängig ist von der Motivation, die dahinter liegt. Doch im Gegensatz zu Hume, der den Menschen in seinen Handlungen durch tugendhafte oder boshafte Charaktereigenschaften motiviert sieht, ist das menschliche Handeln nach Kant von Universalprinzipien motiviert, die sich allein aus der Vernunft ergeben. Begriffe wie *Autonomie* und *Universalität* sind in seiner Moraltheorie von zentraler Bedeutung.

Um zu verstehen, warum die Autonomie in seiner Lehre eine so große Rolle spielt, muss man zunächst verstehen, wie Kant den Menschen in seiner natürlichen Welt sah. Er glaubte, dass das Verhalten von Tieren ausschließlich durch (natürliche) Kräfte, die auf die Tiere einwirken, kausal bestimmt wird. Sie fressen, vermehren sich, töten, ziehen Junge auf und so fort – ein rein instinktives, triebhaftes

Verhalten, das durch physische Ursachen ausgelöst wird. Aus diesem Grund sind moralische Konzeptionen nicht auf Tiere anwendbar. Ein Hund begeht kein Verbrechen, wenn er eine Katze tötet, auch wenn diese Tat in uns starke Emotionen weckt. Tiere sind nicht fähig, vernünftig zu denken oder zu entscheiden, wie sie handeln wollen, und können folglich auch nicht moralisch (und rechtlich) verantwortbar gemacht werden.

Anders der Mensch. Wir sind sehr wohl fähig und imstande, vernünftig zu denken und mithin zu entscheiden, wie wir handeln wollen. Wir können folglich durchaus moralisch (und rechtlich) zur Verantwortung gezogen werden. Und weil wir als Menschen vernünftig denken können, sind wir *autonome* Wesen. Weil wir vernünftig denken können, können wir entscheiden, wie wir handeln wollen; unser Verhalten ist nicht kausal bestimmt – es ist nicht triebhaft oder reflexhaft. Insofern ist Kants Sicht von der Welt klar unterteilt in Menschen als autonome Wesen, die Zweck an sich sind (eigenverantwortlich), und Tiere als nichtautonome Wesen, die sich triebhaft oder reflexhaft verhalten. Der *freie Wille* ist nach Kant die Fähigkeit des Menschen, Entscheidungen treffen zu können.

Doch es gibt eine Lücke in seiner Argumentation: Wenn wir rein rationale, sprich durch Vernunft motivierte Wesen wären, würden wir niemals Fehler machen. Doch wir sind weder rein triebhafte noch rein rationale Wesen. Wir sind irgendwo dazwischen und können deshalb auch falsche Entscheidungen treffen. Manchmal handeln wir gegen die Vernunft und geben unseren Trieben nach, und manchmal handeln wir nach Prinzipien. Oft haben wir die Möglichkeit, unter mehreren Verhaltensweisen zu wählen, die uns

mal mehr, mal weniger gut an unser Ziel führen. Wenn wir beispielsweise Geld brauchen, können wir es einem anderen gewaltsam abnehmen, wir können uns welches leihen oder wir können es redlich verdienen, indem wir eine Ware oder Dienstleistung anbieten. Wir brauchen Regeln und Normen, an denen wir unsere Entscheidungen ausrichten können, die uns sagen, wie wir *entscheiden sollen*, wann immer wir die Entscheidungsmacht haben. Wir brauchen Handlungsregeln, die uns sagen, wie wir uns *verhalten sollen*.

Wie aber finden wir diese Handlungsregeln? Kant war der Überzeugung, dass es ausgeschlossen sei, Regeln (oder Handlungen) in Abhängigkeit von ihren Konsequenzen zu bewerten, da wir die Konsequenzen nicht kontrollieren können. Selbst die besten Entscheidungen können unvorhersehbare verheerende Folgen haben. Was wir aber kontrollieren können, sind unsere Absichten – unsere Motive –, die unseren Handlungen zugrunde liegen. Die Moralität einer Handlung, sprich, ob wir eine Handlung gutheißen oder nicht, hängt also von der ihr zugrunde liegenden Motivation (dem Handlungsgrund) ab. Nach Kant gibt es nur ein Kriterium, welches eine moralische Handlung bedingungslos gutheißt, den guten Willen: Ich beabsichtige, das Richtige zu tun, und richte all meine Handlungen an diesem Prinzip aus. Es ergibt also keinerlei Sinn zu sagen, man habe das Falsche aus den richtigen Gründen getan. Erfolgt eine Handlung aus guten Entscheidungsgründen, so ist sie gut und richtig.

Gleichwohl kann man das Richtige nicht aus den falschen Gründen tun. Kant wäre mit Hume wohl einig darin, dass jemand, der Geld spendet, weil er sich dazu gezwungen sieht oder sich geschäftliche Vorteile erschleichen will,

nicht moralisch gut handelt. Um als moralisch gute und damit richtige Handlung zu gelten, muss die Spende freiwillig erfolgen. Im Gegensatz zu Hume jedoch würde Kant eine Spende als moralisch nicht gut heißen, wenn sie nur dadurch motiviert ist, sich spendabel zu zeigen oder eigennützige Interessen zu verfolgen. Für Kant sind diese Beweggründe rein subjektiv und damit irrelevant. Um moralisch gut zu handeln, kommt es laut Kant einzig darauf an, nach den richtigen Prinzipien zu handeln, und das eigene Handeln in bewusster Absicht daraus zu abzuleiten. Moralisch gutes Handeln ist demnach ein Handeln aus Pflicht.

Wie aber finden wir heraus, was diese *richtigen Prinzipien*, diese Pflichten, sind? Nach Kant finden wir sie durch die reine Vernunft. Oberstes Merkmal eines allgemeinen Moralprinzips ist seine Universalisierbarkeit (Allgemeingültigkeit): Moralität muss für *jedermann* gültig sein – sie lässt keine Ausnahmen zu. Eine Handlungsregel wird erst dann zu einem moralischen Gesetz, wenn sie der Überprüfung nach der obersten, allumfassenden Formel standhält, die da lautet: Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, dass sie ein allgemeines Gesetz werde. Ein universelles Gesetz, so Kants Überzeugung, könne einzig von der Vernunft bestimmt werden, denn die Erkenntnisse der Vernunft seien für jeden vernünftig handelnden Menschen gleich.

Um seine Position zu verdeutlichen: Weil Menschen autonome, vernunftbegabte Wesen sind und weil die Vernunft bei allen vernünftig handelnden Menschen zu denselben Erkenntnissen führt, können wir für uns selbst erkennen, was wir in welcher Situation tun sollen, um moralisch richtig zu handeln. Wir brauchen keinen Staat, keine Religion

oder sonst eine Autorität, die uns sagt, was wir tun sollen. Wir bestimmen selbst, was wir tun. Und wenn wir dieses Tun an den richtigen Prinzipien ausrichten, die wir durch die reine Vernunft erkannt haben, dann ist eine Handlung moralisch – ungeachtet ihrer Konsequenzen.

Kant unterschied zwei allgemeine Gesetze: den hypothetischen und den kategorischen Imperativ. Ein hypothetischer Imperativ ist eine bedingendes Handlungsgebot („wenn"-Gebot), das uns sagt, was wir tun sollen, um ein bestimmtes Ziel zu erreichen – wie beispielsweise: „*Wenn* du willst, dass dir die Menschen vertrauen, mach keine Versprechungen, die du nicht vorhast zu halten." Aber ein solches Gebot taugt nicht als kategorischer Imperativ, als *bedingungsloses* moralisches Gesetz, das universell verpflichtend ist. Es besagt zum einen, „*wenn* du willst, dass dir die Menschen vertrauen", zum anderen aber lässt es sich nicht auf Menschen anwenden, die sich nicht darum scheren, ob irgendwer ihnen vertraut. Und dennoch scheint es moralisch falsch zu sein, Versprechungen zu machen, die man nicht halten kann. Zudem verweist dieser hypothetische Imperativ auf Konsequenzen (Menschen, die einem trauen oder nicht), und diese sind für Kant nicht relevant. Ein kategorischer Imperativ lässt keine Ausnahmen zu, denn die Richtigkeit einer Handlung hängt nicht von subjektiven Wünschen oder von ihren Konsequenzen ab.

Damit eine Handlungsregel (*Maxime*) als kategorischer Imperativ gelten kann, dürfen sich darin keinerlei Widersprüche (*Inkonsistenzen*) finden, was uns die Vernunft ermöglicht zu erkennen. Schließlich kann etwas nicht Kreis und Quadrat gleichzeitig sein. Das wäre ein Widerspruch. „Du sollst die Quadratur des Kreises erreichen" wäre als

kategorischer Imperativ schlicht unmöglich. Um herauszufinden, ob eine Regel den Anforderungen des kategorischen Imperativs genügt und als uneingeschränktes Pflichtgebot gelten kann, ist zunächst die Frage zu stellen, ob ich wollen kann, dass sie für alle uneingeschränkt und ohne Ausnahme Pflicht ist. Erst wenn sich diese Frage widerspruchsfrei beantworten lässt, liegt ein kategorischer Imperativ vor. *Kein* kategorischer Imperativ wäre zum Beispiel diese Regel: „Du hast die Pflicht, Versprechungen zu machen, die du nicht vorhast zu halten." Warum? Überlegen Sie doch einmal: Wenn jeder falsche (lügenhafte) Versprechungen machen dürfte, wären alle Versprechungen sinnlos (denn niemand würde glauben, dass ihm etwas versprochen sei); alle Vorstellungen von *Versprechen* würden in sich zusammenfallen und der Begriff wäre inhaltsleer. Insofern kann diese Regel nicht die Form eines kategorischen Imperativs haben, denn wir können sie nicht zu einem universellen Gesetz erheben, ohne uns in logische Widersprüche zu begeben. Die Regel „Du hast die Pflicht, Versprechen zu halten" hingegen taugt als kategorischer Imperativ, da sie keine Widersprüche in sich birgt. Sie ist zugleich eine Pflicht.

Kant stellt mindestens zwei Formulierungen des kategorischen Imperativs vor (*Grundlegung zur Metaphysik der Sitten*, 1785):

1. „Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde (Grundformel)."[4]

[4] Maio G (2012) Mittelpunkt Mensch: Ethik in der Medizin. Schattauer, Stuttgart, S. 27

2. „Handle so, daß du die Menschheit, sowohl in deiner Person als in der Person eines jeden andern, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchest (Selbstzweckformel)."[5]

Zu den moralischen Pflichten gehören nach Kant die Pflicht, unser eigenes Leben zu bewahren, die Pflicht, wohltätig zu sein, wenn es uns möglich ist, und die Pflicht, das eigene Glück zu sichern. Die zweite Formulierung beinhaltet einen weiteren Grund dafür, warum unsere Reaktionen auf die Standardversion des Trolley-Problems und dessen Variante mit dem dicken Mann so unterschiedlich ausfallen: Die Variante mit dem dicken Mann verlangt von uns, ein menschliches Wesen als Mittel zu einem Zweck zu gebrauchen; und dies verstößt gegen den kategorischen Imperativ, wonach alle Menschen mit Achtung und Respekt zu behandeln sind. Insofern drückt Kant vielleicht explizit aus, wovon wir implizit ausgehen – dass eine solche Handlung ein grundlegendes Menschenrecht verletzt.

Für Hume, so wissen wir, ist das höchste Verdienst, das die menschliche Natur zu erreichen fähig ist, wohlwollende Menschlichkeit. Auch Kant erörtert dieses Wohlwollen:

> „Noch denkt ein *Vierter*, dem es wohl geht, indessen er sieht, dass Andere mit großen Mühseligkeiten zu kämpfen haben, (denen er auch wohl helfen könnte:) was geht's mich an? […] Aber, obgleich es möglich ist, dass nach jener Maxime ein allgemeines Naturgesetz wohl bestehen könnte: so ist es doch unmöglich, zu *wollen*, dass ein solches Princip als Naturgesetz allenthalben gelte. Denn ein Wille, der dieses beschlösse, würde sich selbst widerstreiten,

[5] ebenda

> indem der Fälle sich doch manche ereignen können, wo er Anderer Liebe und Theilnehmung bedarf, und wo er, durch ein solches aus seinem eigenen Willen entsprungenes Naturgesetz sich selbst alle Hoffnung des Beistandes, den er sich wünscht, rauben würde."[6]

Egoismus taugt nach Kant demnach nicht als ein kategorischer Imperativ, denn Egoismus verhindert, dass wir Hilfe bekommen, wenn wir sie brauchen. Dies erklärt einmal mehr, warum wir uns so schwer tun mit dem Trolley-Problem. Aus rein egoistischer Sicht könnten wir fragen: „Was geht's mich an, ob einer stirbt oder fünf?" Doch wohlfeiles Handeln stellt eine Pflicht dar, und diesem kategorischen Imperativ nicht nachzukommen, einen Widerspruch.

Es gibt zwei weitere deontologische Prinzipien, die Kant nicht betrachtet, die aber dennoch eine kurze Erörterung verdienen. Zum einen wäre da die *Doktrin vom Tun und Unterlassen* (Doktrin vom aktiven und passiven Tun). Sie besagt, dass es durchaus Bedingungen geben kann, unter denen die Verursachung eines Übels moralisch erlaubt ist. So darf der Handelnde die (moralisch) schlechte Folge nicht aktiv herbeiführen, darf sie aber zulassen: Angenommen, der dicke Mann befände sich bereits auf den Gleisen, und Sie täten nichts, um ihn zu retten, weil Sie dadurch fünf andere Menschenleben retten. Nun könnte man sagen, es sei ein minder schweres Vergehen, den Mann nicht zu retten, da Sie das Unglück durch passives Tun schlicht geschehen lassen und den dicken Mann nicht durch aktives Tun auf die Gleise stoßen. Das gleiche Argument gilt aber auch für die Straßenbahn; das deontologische Prinzip näm-

[6] Kant I (1785) Grundlegung zur Metaphysik der Sitten; zitiert nach: Hartenstein G (Hrsg) (1867) Immanual Kants's sämmtliche Werke, Bd. 4, S. 271

lich legt fest, dass es grundsätzlich falsch ist, einen unschuldigen Menschen zu töten, und damit ist es auch falsch, die Weiche umzustellen – das Entscheidungsdilemma ist also vorprogrammiert.

Das zweite deontologische Prinzip, das Beachtung verdient, ist das *Prinzip der Doppelwirkung*. Danach macht es einen erheblichen Unterschied, ob der Handelnde die (moralisch) schlechte Folge seiner Handlung beabsichtigt, oder ob er sie bloß vorhersieht und sie eine lediglich unbeabsichtigte Nebenwirkung darstellt. Betrachten wir noch einmal die Standardversion des Trolley-Problems und dessen Variante mit dem dicken Mann: Wenn Sie den dicken Mann auf die Gleise stoßen, verstoßen Sie gegen dieses Prinzip, da Sie beabsichtigen, ihn zu töten, auch wenn Sie zugleich beabsichtigen, andere Menschen zu retten. Ihn zu töten, ist integraler Bestandteil Ihres Vorhabens, andere zu retten. In der Standardversion ist Ihre Absicht, die Weiche umzustellen, um Menschen zu retten, nicht, die einzelne Person auf dem Gleis zu töten. Diese Person zu töten, ist nicht notwendiger Bestandteil Ihres Vorhabens, jedoch eine unbeabsichtigte Nebenwirkung. Die Variante mit dem dicken Mann untergräbt also das Prinzip der Doppelwirkung, die Doktrin vom Tun und Unterlassen sowie auch die zweite Formulierung des kategorischen Imperativs. Kein Wunder, dass wir die Frage, ob es moralisch statthaft sei, den dicken Mann zu töten, mit „Nein" beantworten.

Ich möchte diesen Abschnitt über Kant nicht beschließen, ohne noch kurz den „Mörder an der Tür" zu betrachten. Seine Fokussierung auf den kategorischen Imperativ führt Kant bisweilen zu Positionen, die den meisten von uns, na, sagen wir mal, etwas verrückt erscheinen. Seine

Antwort auf Benjamin Constants Dilemma vom „Mörder an der Tür" ist so ein Fall. Das Dilemma ist folgendes:

> Jemand bittet Sie, ihn in Ihrem Haus zu verstecken, weil irgendwer hinter ihm her ist, um ihn zu ermorden. Sie kommen seiner Bitte nach und verstecken ihn. Dann kommt der vermeintliche Mörder an Ihre Tür und will wissen, ob sich die Person im Haus befindet. Sagen Sie ihm die Wahrheit oder lügen Sie ihn an?

In seiner Antwort auf dieses Dilemma hält er eisern am kategorischen Imperativ fest, der es grundsätzlich verbietet, zu lügen, auch unter diesen Umständen. Die aus der Vernunft geborene Pflicht gebietet uns, die Wahrheit zu sagen und das Leben dieser Person aufs Spiel zu setzen. Wir erinnern uns: Für Kant sind Konsequenzen irrelevant. Solange wir nach vernunftbasierten Prinzipien handeln, tun wir das (moralisch) Richtige. Die Konsequenzen haben wir selbst nicht in der Hand. Die meisten Menschen aber lehnen diese rigide Anwendung des Pflichtverhaltens ab, da wir die Konsequenzen unseres Handelns – in diesem Fall den möglichen Tod eines Menschen – nicht grundsätzlich ignorieren können.

Was Jeremy Bentham und John Stuart Mill zu sagen haben

Jeremy Bentham (1748–1832) war ein englischer Philosoph und Jurist, der in seinem Werk *An introduction to the principles of morals and legislation* (1789) eine Moraltheorie entwickelte, die als Utilitarismus bekannt ist. John Stuart

Mill (1806–1873), ebenfalls englischer Philosoph und Ökonom, griff diese Theorie auf und erweiterte sie. Zu seinen wichtigsten Werken zählen *Über die Freiheit* (1859), in dem er die Bedeutung der Individualität betont; *Der Utilitarismus* (1861), in dem er Benthams Theorie erweitert, sowie *Die Hörigkeit der Frau* (1869), in dem er die Rechte der Frauen hochhält. (1865 zieht Mill für die liberale Partei der Whigs ins Parlament ein und ist der erste Parlamentarier, die für das Wahlrecht von Frauen kämpft.) Der Utilitarismus ist ein konsequenzialistisches Prinzip, das eine Handlung oder eine Politik als richtig bewertet, wenn sie „das größte Glück für die größte Zahl von Menschen" hervorbringt. Dieses Prinzip vom „größten Glück" ist Grundformel und Maxime des Nützlichkeitsprinzips.

Auch wenn es Hume war, der den Gedanken der Nützlichkeit erstmals in die Moraltheorie einbrachte, ist es Bentham, der den Maßstab der moralischen Richtigkeit einer Handlung an erlebte Freude und erlebten Schmerz bindet. Im Vorwort zu *An introduction to the principles of morals and legislation* schreibt er:

> „Die Natur hat die Menschheit unter die Herrschaft zweier souveräner Gebieter – *Leid* und *Freude* – gestellt. Es ist an ihnen allein aufzuzeigen, was wir tun sollen, wie auch zu bestimmen, was wir tun werden. Sowohl der Maßstab für Richtig und Falsch als auch die Kette der Ursachen und Wirkungen sind an ihrem Thron festgemacht. Sie beherrschen uns in allem, was wir tun, was wir sagen was wir denken."[7]

[7] Bentham J (2013) Eine Einführung in die Prinzipien der Moral und Gesetzgebung; zitiert nach: Höffe O (2008) Einführung in die utilitaristische Ethik. Francke, Tübingen, S. 55

Benthams Nutzenprinzip (auch als das „Prinzip des größten Glücks“ bezeichnet) lässt sich wie folgt zusammenfassen: Das einzige Gute ist die Freude, das einzig Schlechte ist der Schmerz. Um den moralischen Wert einer Handlung zu beurteilen, müssen wir also fragen, ob sie dem Empfänger Freude oder Schmerz bereitet. Nach diesem Prinzip sind alle Handlungen geboten, die Freude und Glück bringen, und alle Handlungen verboten, die Schmerz und Leid verursachen. Moralisch zu handeln, heißt also, Freude und Glück zu fördern und Schmerz und Leid zu vermeiden bei denjenigen, deren Interessen durch meine Handlung betroffen sind. Mill greift Benthams Nutzenprinzip als Grundlage für seine eigene Theorie auf, die er wie folgt definiert: *Handlungen sind insoweit und in dem Maße moralisch richtig sind, als sie die Tendenz haben, Glück zu befördern, und insoweit moralisch falsch, als sie die Tendenz haben, das Gegenteil von Glück zu bewirken*. Der Utilitarismus von Mill und Bentham ist eine Variante des Konsequenzialismus, ein Zweig der Philosophie, der die Moralität einer Handlung nach ihren Konsequenzen bemisst.

Sowohl Bentham als auch Mill beziehen ihr moralisches Konzept auf empirisch feststellbare Konsequenzen, und zwar weitgehend deshalb, weil sie sich damit bewusst abheben wollen von den intuitiven Ethiken. Der sogenannte Intuitionismus postuliert, dass es objektive moralische Wahrheiten gibt, die dem Verstand durch eine intuitive Wissbarkeit unmittelbar einsichtig sind. Bentham und Mill stemmen sich vehement gegen dieses Prinzip der Moral, denn wenn die Richtigkeit moralischer Regeln intuitiv wissbar sein kann, so argumentieren sie, dann wären sie *unanfechtbar*. Moralische Werte ließen sich einfach endlos

behaupten und immer wieder bekräftigen, und zwar ohne irgendeine Angabe von Gründen. Dem Despotismus wären damit Tür und Tor geöffnet; denn wenn ich ausgestattet bin mit der nötigen Macht, dann regiert allein meine intuitive Wahrheit, ohne Rücksicht der Konsequenzen für andere. Das Gleiche gilt für Kants ethischen Rationalismus (der die Vernunfteinsicht als Prinzip moralischen Handelns postuliert). Mill wirft Kant vor, die Notwendigkeit moralischer Gesetze (die er mit seiner Formulierung des kategorischen Imperativs als gegeben ansieht) nicht ausreichend zu begründen, wenn diese zu Widersprüchen führen. Stattdessen, so betont Mill, weise er moralische Gesetze zurück, sobald sie zu unerwünschten Folgen führen.

Und um noch einmal auf unser Trolley-Problem zurückzukommen: Ein strenger Utilitarist würde sich dafür entscheiden, die Weiche umzustellen, beziehungsweise den dicken Mann in den Tod zu stoßen.

Was uns richtig erscheint: die Psychologie der moralischen Urteilsbildung

Es gibt zahllose Studien, in denen die Teilnehmer hypothetische Dilemmas wie das Trolley-Problem vorgelegt bekamen. Wie sich zeigt, wägen die Teilnehmer dabei keineswegs nur die Anzahl der geretteten gegen die Anzahl der getöteten Menschen ab. In Dutzenden von Studien sagen 80 % der Teilnehmer „Ja“ zu der Option, die Weiche umzustellen, und „Nein“ zu der Option, den Mann auf die

Gleise zu stoßen. Im folgenden Dilemma aber fielen die Antworten tendenziell ungefähr halbe-halbe aus (Greene et al. 2001):

> Feindliche Soldaten haben Ihr Dorf eingenommen und sollen laut Befehl alle verbliebenen Zivilisten töten. Sie haben sich mit einigen anderen Dorfbewohnern in den Keller eines großen Hauses geflüchtet. Oben hören Sie die Stimmen von Soldaten, die das Haus nach Wertsachen durchsuchen. Ihr Baby beginnt laut zu schreien, daher halten Sie ihm den Mund zu, damit es Sie nicht verrät. Wenn Sie die Hand wegnehmen, wird sein Weinen die Aufmerksamkeit der Soldaten erregen, die Sie, Ihr Kind und die anderen im Keller versteckten Menschen umbringen werden. Um sich selbst und die anderen zu retten, müssen Sie Ihr Kind ersticken. Halten Sie es für richtig, Ihr Kind zu ersticken, um sich und die anderen Dorfbewohner zu retten?

Auch hier müssen Sie sich entscheiden, ob Sie ein Leben opfern, um viele andere zu retten – genau wie im Trolley-Problem. Und auch hier scheinen wir keineswegs nur die Anzahl der geretteten gegen die Anzahl der getöteten Menschen abzuwägen.

Unzählige Studien wurden durchgeführt, um zu untersuchen, woran die Teilnehmer ihre Beurteilungen der verschiedenen moralischen Dilemmas festmachen. In manchen (wie beim Trolley-Problem oder beim oben beschriebenen Problem) ging es darum, Leben zu retten oder zu opfern. In anderen ging es um weniger fatale Umstände, etwa darum, den Lebenslauf zu fälschen, um einen begehrten Job zu bekommen, oder den Partner zu betrügen. Wie sich eine Person entscheiden wird, kann anhand ihrer indi-

viduellen Entscheidungspräferenzen vorhergesagt werden, aber auch anhand von Entscheidungsgründen, denen sie sich moralisch verpflichtet sieht. Teilnehmer, die vornehmlich ihrer Intuition folgen, beantworten die Schlussfrage im obigen Dilemma tendenziell mit „Nein", während diejenigen, die bewussten Überlegungen folgen, sie tendenziell mit „Ja" beantworten (Baumert 2010). Das heißt, dass all die, die lieber gründlich überlegen, bevor sie eine Entscheidung fällen, meist utilitaristische Entscheidungen treffen. Es gibt also zwei Lager: Wer seine Moralphilosophie nach dem utilitaristischen Prinzip ausrichtet, beantwortet die Schlussfrage also tendenziell mit „Ja", während derjenige, der es mit dem deontologischen Prinzip hält, sie tendenziell mit „Nein" beantwortet.

Das mag Sie nun nicht wirklich überraschen, aber vielleicht staunen sie ja über folgende Ergebnisse: Wurde die Zeit, innerhalb der die Teilnehmer ihre Entscheidung treffen mussten, verringert, gab es weniger utilitaristische Entscheidungen. Zum Beispiel fiel der Anteil der Personen, die sich für das Umstellen der Weiche entschieden, von 80 auf 70 %; und der Anteil der Personen, die sich für das Ersticken des Babys entschieden, fiel von 45 auf 13 %. Es braucht offenbar mehr Entscheidungszeit, um diese Handlungen für moralisch gut zu befinden (Cummins 2011).

Noch erstaunlicher aber ist der Einfluss von scheinbar irrelevanten Faktoren auf die moralische Beurteilung einer Handlung. Vergleichsweise weniger Teilnehmer waren willens, die Weiche umzulegen, wenn ihnen das Dilemma mit dem dicken Mann zuerst vorgelegt worden war. Und im umgekehrten Fall zeigte sich, dass vergleichsweise mehr Teilnehmer willens waren, den dicken Mann in den Tod

zu stoßen, wenn man sie zuerst mit der Standardversion des Trolley-Problems konfrontiert hatte. Das bedeutet, dass moralische Urteile variieren, je nachdem, wie die Teilnehmer dachten, bevor sie ein Urteil fällten. Und das gilt nicht nur für ganz gewöhnliche Leute, professionelle Philosophen verhalten sich nicht anders (Schwitzgebel & Cushman 2012). (Da fragt man sich, ob richterliche Urteile variieren, je nachdem, welche Art von Fall der Richter vor seiner Entscheidung gerade gehört hat.) Und ebenfalls erstaunlich sind die Auswirkungen von Faktoren, die belangloser scheinen als richterliche Urteile: Die Teilnehmer beurteilen eine Handlung tendenziell eher als moralisch falsch, wenn sie ihre Entscheidung im Rahmen eines ekelerregenden Umfelds treffen (Schnall et al. 2008a, b) – wie beispielsweise neben einem gebrauchten Taschentuch zu sitzen!

Für die Tatsache, dass diese Faktoren einen so großen Einfluss auf die moralische Urteilsbildung haben, gibt es zwei Erklärungen, die einmal mehr einen sorgfältigen Unterschied zwischen zwei grundsätzlich verschiedenen Denksystemen machen. Die erste Erklärung liefert der US-amerikanische Psychologe Jonathan Haidt. Die moralischen Intuitionen, so sagt er (genau wie Hume), kommen zuerst, die bewussten Überlegungen folgen. Dieser bewusste, räsonierende Teil, so Haidt weiter, diene vor allem dazu, die Entscheidungen und Neigungen des unbewussten, intuitiven Teils im Nachhinein rational zu begründen und zu rechtfertigen (Haidt 2007). Die zweite Erklärung liefert der Neurowissenschaftler Joshua Greene mit seiner sogenannten Zwei-Prozess-Theorie, die moralische Urteile als das Ergebnis eines Konkurrenzkampfs zwischen emotional-intuitiven und analytisch-vernunftorientierten Prozessen be-

schreibt (Greene 2007), die der moralischen Urteilsfindung zugrunde liegen.

Greenes Zwei-Prozess-Theorie steht ganz im Einklang mit Humes Sicht der moralischen Urteilsbildung. Das emotional-intuitive Denken ist Kernkompetenz von System 1, das schnelle, automatische und für gewöhnlich emotional gesteuerte Urteile/Entscheidungen liefert. Die Beurteilung einer Handlung oder des Charakters einer Person ist in ihrer Natur nichts weiter als eine simple Bewertung aus dem Bauch heraus nach dem Schema Gut-Böse oder Gefallen-Missfallen. Urteile dieser Art kommen in unserem Bewusstsein an, ohne dass wir die einzelnen Phasen der Urteilsbildung bewusst wahrgenommen hätten – die Suche nach brauchbaren Hinweisen, das Abwägen von Erkenntnissen oder gar das Ableiten einer Schlussfolgerung.

Das analytisch-vernunftorientierte Denken ist Kernkompetenz von System 2. Nach Haidt setzt es erst nach der intuitiven Beurteilung ein, um eine gefasste Entscheidung nachträglich zu rechtfertigen und rational zu begründen. Es ist ein kontrollierter, weniger emotionaler Prozess, der bedingt ist durch bewusste gedankliche Aktivität. Er umfasst die Verarbeitung von Informationen über Menschen und ihre Handlungen, um zu einem moralischen Urteil oder einer Entscheidung zu gelangen. Für Haidt ist ein moralisches Urteil daher eine Art ästhetisches Urteil: Wir sehen eine Handlung oder hören eine Geschichte und haben ein spontanes Gefühl dafür, ob uns das, was wir sehen oder hören, gefällt oder nicht. Wir können nicht erklären warum, aber wir können uns eine vernünftige Rechtfertigung für unser bereits gefasstes Urteil zurechtlegen.

Um seine Theorie zu stützen, zitiert Haidt das Phänomen vom gespaltenen Geist. Wenn der Geist gespalten ist, so Haidt, wissen wir intuitiv, wann eine Handlung moralisch falsch ist, auch wenn wir nicht erklären können, warum. Stellen wir uns zum Beispiel einmal vor, zwei Geschwister, Bruder und Schwester, haben ein einziges Mal miteinander geschlafen; niemand sonst weiß davon, die beiden haben verhütet, keiner der beiden hat Schaden davon getragen, und beide haben das Gefühl, als Geschwister seither einander näher zu sein. Fragt man nun einzelne Leute, ob diese Handlung moralisch statthaft sei, würde die Mehrheit der Befragten mit „Nein" antworten. Aber warum sie mit „Nein" antworten, können sie nicht erklären.

Auch der Neurowissenschaftler Joshua Greene beruft sich in seiner Erklärung auf duale Prozesse (wie sie auch Kahneman und andere beschreiben), wonach Intuition und Verstand heftig miteinander konkurrieren, um die Vorherrschaft zu erlangen, auf der dann die endgültige Entscheidung basiert. Fallen die Ergebnisse dieser beiden Denksysteme gleich aus, erfolgt die endgültige Entscheidung leicht und schnell. Fallen sie jedoch unterschiedlich aus, ist der Konflikt nicht gelöst. Und das macht die endgültige Entscheidung langsam und schwierig, zumal jeder Prozess den anderen jederzeit außer Kraft setzen kann, um den Sieg über die endgültige Entscheidung zu erringen. Man fühlt sich hin- und hergerissen, als schlügen zwei Seelen in einer Brust. Die Vernunft braucht Zeit, um ein klares Bild zu liefern, das jedoch der anfänglichen emotional-intuitiven Reaktion nicht selten zuwiderläuft. Die endgültige Entscheidung wird schlicht und einfach die sein, die das Siegersystem durchsetzt – welches auch immer es ist.

Greene stützt seine Theorie auf zahlreiche Studien, in denen die Probanden Entscheidungen über moralische Dilemmas trafen, während sie an ein MRT-Gerät angeschlossen waren, das den Forschern Bilder ihres Gehirns lieferte (Greene et al. 2001, 2004). Was die Neurowissenschaftler als erstes bemerkten, war, dass die moralische Urteilsbildung direkt gekoppelt ist an die menschliche Fähigkeit, Mitgefühl oder Emotionen zu empfinden (wie bereits Hume theorisierte). Zum Beispiel sind die Hirnareale, die aktiv sind, wenn man selbst Schmerz empfindet, dieselben, die aktiv sind, wenn ein normal entwickeltes Kind oder ein normal entwickelter Erwachsener einen anderen sieht, der Schmerz empfindet. Bei Schmerzempfindungen sind zwar mehrere Hirnareale aktiv, das allerwichtigste für die moralische Empathie aber ist der ventromediale präfrontale Cortex (VMPC).

In einer Studie von Heckeren et al. (2003) sollten die Teilnehmer Aussagen lesen, wie „A stiehlt ein Auto“ oder „A schwärmt für ein Auto“, und ein Urteil darüber fällen, ob die jeweiligen Aussagen moralisch vertretbar seien. Daneben bekamen sie nichtmoralische Aussagen zu lesen, wie „ A geht spazieren“ und „A geht jeden Moment spazieren“, über deren inhaltliche Aussage sie ebenfalls ein moralisches Urteil fällen sollten. Wie sich herausstellte, war der VMPC sehr viel aktiver, wenn es moralische Aussagen zu befinden galt, als wenn es um nichtmoralische ging. Moll et al. (2001) erhielten ähnliche Ergebnisse, als sie im Rahmen einer Studie die Probanden aufforderten, Entscheidungen darüber zu fällen, ob eine einfache moralische Aussage (z. B. „Wenn es sein muss, verstoßen wir gegen das Gesetz“) sowie eine einfache nichtmoralische Aussage (z. B. „Steine bestehen

aus Wasser") richtig ist oder falsch. Und auch hier zeigte sich, dass der VMPC besonders aktiv war, wenn es um die Verarbeitung moralischer Aussagen ging. Der VMPC wird also aktiviert, so die wissenschaftliche Interpretation, wenn wir auf sozio-emotionale Aspekte einer moralischen Situation reagieren.

Greene et al. (2001, 2004) gingen einen Schritt weiter, indem sie den Probanden nicht nur moralische Aussagen zur Beurteilung vorlegten, sondern auch klassisch moralische Dilemmas wie das Trolley-Problem oder das Problem mit dem schreienden Baby, und mittels bildgebender Verfahren beobachteten, welche Hirnareale während der Urteilsbildung besonders aktiv waren. Auch hier stellte sich heraus, dass der VMPC sehr viel aktiver war, wenn es um moralische und nicht um nichtmoralische Dilemmas ging, was darauf deutet, dass die Teilnehmer moralische Dilemmas emotional sehr viel intensiver erlebten. Wenn es hingegen darum ging, ein utilitaristisches Urteil zu fällen, so waren die Hirnareale, die in kognitive Konfliktsituationen (anteriorer cingulärer Cortex, ACC) sowie in abstrakt-logische Denkvorgänge (dorsolateraler präfrontaler Cortex, dlPFC) involviert sind, hochgradig aktiv. Diese Ergebnismuster zeigten, so schließt Greene, dass Emotion und Verstand während moralischer Urteilsbildungsprozesse neurologisch trennbar sind: Emotionale Reaktionen im VMPC treten zuerst auf; erst danach wird ein anderes Hirnareal aktiviert (dlPFC), um sich über diese anfängliche emotional-intuitive Reaktion hinwegzusetzen und am Ende zu einem utilitaristischen Urteil zu gelangen – zu einem wohlüberlegten Urteil also, das im moralischen Sinne gut bzw. richtig ist.

Koenigs et al. (2007) verglichen die moralischen Urteile von Patienten, die Schädigungen des VMPC aufwiesen, mit denen von Patienten mit Schädigungen anderer Hirnareale und denen einer gesunden, nicht hirngeschädigten Vergleichsgruppe, gegliedert nach Alter, Geschlecht und weiteren Faktoren. Dabei stellten sie fest, dass die VMPC-Patienten zu einer Reihe moralischer Dilemmas erheblich mehr utilitaristische Urteile abgaben. Woran dies liegt, erklärt die Wissenschaft zum einen damit, dass wir sehr viel eher vernunftorientierte Urteile treffen, wenn Emotionen außen vor bleiben. Zum anderen sehen die Forscher darin einen Hinweis, dass Hirnschädigungen offenbar streng utilitaristische Urteile befördern. So oder so, wie es scheint, sind Emotion und Verstand tatsächlich neurologisch separierbar und können durch organische Hirnschädigungen dissoziiert werden.

Und wie steht es mit Kants deontologischen Betrachtungen? Vergleichbare Studien aus der Verhaltenspsychologie und der Neurowissenschaft kamen ebenfalls zu dem Ergebnis, dass moralische Urteile stark von emotionalen Faktoren abhängen (Borg et al. 2006). Die Probanden beurteilten eine Handlung, die beabsichtigte negative Folgen hat, als moralisch weniger statthaft, als wenn die Folgen unbeabsichtigt sind. Und sie beurteilten eine Handlung, die ergriffen wird und negative Folgen hat, als moralisch weniger statthaft, als wenn sie unterlassen wird, auch dann, wenn die (negativen) Folgen daraus dieselben sind (z. B. wenn fünf Menschen so oder so den Tod finden). Dies sind aber keine rein analytisch-vernunftorientierten Urteile, die im Widerspruch stünden mit der deontologischen Sicht (wonach bestimmte Handlungen in sich richtig oder falsch

sind, unabhängig von ihren Konsequenzen). Dilemmas, die gegen das Prinzip der Doppelwirkung verstoßen und verbunden sind mit beabsichtigten oder unbeabsichtigten negativen Folgen, lösen intensive Aktivitäten im emotionsorientierten System 1 aus (VMPC). Dilemmas, die gegen die Doktrin vom Tun und Unterlassen verstoßen, aktivieren System 2 vornehmlich nur dann, wenn die (moralisch negativen) Folgen aus einer Handlung dieselben sind, egal ob die Handlung durchgeführt oder unterlassen wird. Wenn die Folgen unterschiedlich sind, findet eine stärkere Aktivität in System 1 statt (VMPC). Es sieht also ganz danach aus, als habe Kant Recht, wenn er sagt, dass moralische Imperative moralischen Urteilen zugrunde liegen und dass wir Menschen keine rein rational denkenden Wesen sind.

Im Rahmen einer weiteren Studie (Young et al. 2010) nutzten die Forscher die transkranielle Magnetstimulation (TMS), um leichte Stromstöße in einen Teil des Gehirns zu induzieren, der besonders wichtig ist für unsere Fähigkeit, mit anderen zu fühlen und uns in sie hineinzuversetzen – der sogenannte rechte temporo-parietale Übergang (der Verbindungsbereich von Temporal- und Parietallappen). Die Stromstöße bewirken, dass dieser Teil des Gehirns vorübergehend außer Gefecht gesetzt und funktionsuntüchtig wird – ungefähr so, als würde der helle Blitz eines Fotoapparats die Sehzellen auf der Netzhaut kurzzeitig ausschalten und dadurch einen blinden Fleck im Sehfeld schaffen. Als die Forscher nun mithilfe der transkraniellen Magnetstimulation den rechten temporo-parietalen Übergang bei ihren Probanden ausschalteten, zeigte sich, dass diese fortan nicht mehr fähig waren, sich in eine andere Person und deren Handlungsabsicht hineinzuversetzen – sie hatten ihr

moralisches Urteilsvermögen verloren. Ein Beispiel: Die Probanden waren aufgefordert, verschiedene Szenerien moralisch zu beurteilen. In einem dieser Szenarien süßt eine Frau namens Grace den Tee ihrer Freundin mit einem weißen Pulver, das sie für Zucker hält. In einem anderen hält sie das weiße Pulver für Gift und gibt es in den Tee ihrer Freundin. In einigen der Szenarien stirbt die Freundin, in anderen überlebt sie. Die meisten von uns würden ihr Urteil davon abhängig machen, ob Grace wissen konnte, was sie tat. Hat sie das Gift versehentlich für Zucker gehalten, so hat sie moralisch nichts Unrechtes getan, auch wenn ihre Freundin dadurch starb. Wusste sie aber, dass es Gift war, so hat sie sich moralisch verwerflich verhalten, auch wenn ihre Freundin überlebte. Doch die Probanden, die mittels TMS leichten Stromstößen ausgesetzt waren, machten ihr moralisches Urteil über das Verhalten von Grace größtenteils davon abhängig, ob die Freundin starb oder nicht. Graces Handlungsabsicht floss in dieses Urteil kaum mit ein.

Wozu überhaupt Moral?

Moralische Fragen erregen unsere Aufmerksamkeit, bewegen nachweislich das Gehirn und lösen starke emotionale Reaktionen aus. Man sollte also meinen, dass Moral und Moralität für unser ganzes Denken, Fühlen und Handeln eine überaus wichtige Funktion erfüllt. Mit Bezug auf das Werk des französischen Philosophen und Soziologen Emile Durkheim (1858–1917) stellt Haidt heraus, dass Moral eine wichtige soziale Funktion erfüllt: Moral bindet und stärkt – sie beschränkt den Einzelnen auf seine Funktion in

einer bestimmten sozialen Gemeinschaft, die eine emergente Entität mit einer eigenen moralischen Ordnung ist. Eine moralische Gemeinschaft hat gemeinsam geteilte Normen und Werte, die Verhaltensregeln vorgeben und das Zusammenleben ordnen, verbunden mit Mitteln und Maßnahmen, Abweichler zu bestrafen und/oder Mitwirkenden Belohnungen zuteilwerden zu lassen.

Insofern erinnert uns die Moral an unsere Diskussion um Kosten und Nutzen für kooperierende und defektierende Spieler; und sie erklärt weitgehend die evolutionären Ursprünge für konkrete Entscheidungen in modellierten Spielsituationen, die häufig von spieltheoretischen Analysen abweichen. Nach Haidt jedoch gibt es weitere Aspekte der Moralität, die ebenso tiefe evolutionäre Wurzeln haben. Basierend auf interkulturellen Studien legen Haidt und Joseph (2008) *Die Theorie der moralischen Grundlagen* vor, wonach sich fünf psychologische Grundlagen für die moralische Urteilsbildung ausmachen lassen, die alle einen eigenen evolutionären Ursprung haben. Hinter dieser Theorie steht die Frage, warum moralische Prinzipien quer durch die Kulturen so stark variieren, trotzdem aber so viele Ähnlichkeiten und wiederkehrende Themen aufweisen. Haidt erklärt dies im Kern seiner Theorie damit, dass es fünf angeborene und universell vorhandene moralische Systeme gibt, die vor jeder Erfahrung existieren und die Grundlage der „intuitiven Moral" bilden.[8] In dieses kognitive Moralbewusstsein hinein weben die einzelnen Kulturen bestimmte Tugenden, Erzählungen, soziale Gebilde und Institutionen

[8] Prof. Jonathan Haidt im Spiegel (2/2013) Wir reiten auf einem Elefanten. http://www.spiegel.de/spiegel/print/d-90438239.html (Zugriff 10.1.2014)

und schaffen dadurch eine kulturelle Vielfalt moralischer Werte und Vorschriften. Die fünf Säulen der Moral sind:

1. Fürsorge – wurzelt in emotionalen Bindungssystemen, will heißen, in der Notwendigkeit, unseren Nachwuchs zu beschützen und zu umhegen. Diese Systeme sind unentbehrlich für unsere Fähigkeit, uns in andere einfühlen und Schmerzen nachempfinden zu können. Damit verbunden sind Tugenden wie Güte, Sanftmut und wohlwollende Hege.
2. Fairness – wurzelt im Prinzip der Reziprozität, einem Altruismus, der auf Gegenseitigkeit beruht (s. Kapitel 2). Diese moralische Grundlage erzeugt Gedanken der Gerechtigkeit, Rechtsgesetze und Autonomie.
3. Loyalität (Gruppenidentifikation) – wurzelt in unserer langen Geschichte als Stammesvölker, die wechselnde Koalitionen bildeten. Damit verbunden sind Tugenden wie Patriotismus und Selbstaufopferung für die Gruppe; es gilt „einer für alle, alle für einen".
4. Autorität – wurzelt in unserer langen Geschichte als Primaten mit einer hierarchisch geordneten Sozialstruktur. Damit verbunden sind Tugenden wie Führerschaft, Gefolgschaft, sich rechtmäßigen Autoritäten zu fügen und Traditionen zu achten.
5. Reinheit – wurzelt in Ekel- und Abscheugefühlen (vor verunreinigten Lebensmitteln etwa), die das menschliche Überleben sichern und vor Krankheiten schützen. Damit verbunden sind religiöse Vorstellungen und Tugenden, das Trachten, edle Pfade zu beschreiten und fleischliche Lüste zu bezähmen.

Nach Haidt basieren die moralischen Strukturen überall auf der Welt auf diesen Grundlagen, jedoch sind sie je nach Kultur unterschiedlich akzentuiert. Interessanterweise stellte sich heraus, dass das moralische Wertesystem der zivilisierten Kulturen der westlichen Welt, die vornehmlich Fürsorge und Fairness in den Mittelpunkt ihrer gesellschaftlichen Strukturen stellen, sehr viel enger gestrickt ist, als das anderer Kulturkreise. Weitaus interessanter jedoch ist, dass das Spektrum moralischer Werte je nach politischer Orientierung offenbar ganz erheblich differiert (Haidt & Graham 2007). Insgesamt zeigte sich, dass sich liberal orientierte politische Systeme in ihren moralischen Urteilen über Handlungen und Maßnahmen fast ausschließlich von Prinzipien der Fürsorge und Fairness leiten lassen. In konservativen Systemen dagegen sind alle fünf Grundlagen nahezu gleichmäßig verteilt. Verstöße gegen Autoritäten oder die Reinheit wiegen genauso schwer wie Verstöße gegen Fairness oder Fürsorge. Dass Konservative und Liberale in wichtigen Fragen der Sozialpolitik häufig keinen gemeinsamen Nenner finden, ist nicht zuletzt auf diese Unterschiede der moralischen Ansätze und Akzente zurückzuführen. Sie blicken durch eine völlig unterschiedliche moralische Brille. Der Konservative kann nicht verstehen, warum der Liberale nicht einsehen mag, warum es moralisch falsch ist, sagen wir mal, die amerikanische Flagge zu verbrennen. Für ihn ist die Flagge heilig, sie zu verbrennen ein Verstoß gegen Autorität und Loyalität. Der Liberale aber hat diese Werte erst gar nicht auf dem Radar.

5
Das Spiel der Logik

Nehmen wir einmal an, in einem etwas weiter entfernten Kino liefe ein Film, den Sie unbedingt sehen wollen. Zu Fuß dorthin zu gehen, wäre zu weit, und ein Auto haben Sie nicht. Ihre Überlegungen könnten in etwa so aussehen:

> Ich will den Film sehen.
> Aber ich habe kein Auto,
> Mein Freund hat ein Auto und er hat vor, sich den Film anzusehen.
> Wenn er mich mitnimmt, ist mein Problem gelöst.
> Ich werde ihn mal anrufen und fragen, ob er mich mitnimmt.

Endergebnis dieser Gedankenkette ist eine Handlung – den Freund anrufen und ihn um eine Mitfahrgelegenheit bitten. Diese Art des Denkens wird als *praktische Vernunft* (oder *methodische* Denkweise) bezeichnet. Sie ist auf eine Handlung ausgerichtet.

Nun hat es der menschliche Geist aber an sich, dass er in einem fort denkt – selbst wenn es gerade gar keine Probleme zu lösen gilt. In den meisten Fällen münden diese Gedanken aber nicht in konkrete Handlungen, sondern

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_5,

in Erkenntnisse. Diese Art des Denkens wird als *theoretische Vernunft* (oder *diskursive* Denkweise) bezeichnet. *Wir nutzen die theoretische Vernunft, um zu entscheiden, welche Erkenntnisse logisch aus anderen Erkenntnissen folgen.* Angenommen, Sie fragen Ihren Freund, ob er Sie ins Kino mitnehmen kann, doch er sagt Ihnen, sein Auto sei kaputt. Daraufhin nimmt ein anderer Freund Sie mit. Direkt vor dem Kino sehen Sie ersteren mit seinem Auto auf den Parkplatz fahren. Sie könnten folgende Gedanken haben:

> Er hat mir erzählt, sein Auto sei kaputt.
> Dabei parkt er gerade ein.
> Wenn sein Auto kaputt wäre, hätte er gar nicht herfahren können.
> Er hat es entweder prompt repariert oder er hat mich angelogen.
> Aber zum Reparieren blieb ja kaum Zeit.
> Also hat er mich angelogen.
> Wenn er mich anlügt, ist er kein richtiger Freund.
> Genau. Er ist kein richtiger Freund.

Diese Gedanken sind nicht zweckhaft oder zielgerichtet. Sie sind lediglich eine Aneinanderreihung von Schlüssen, die in eine Erkenntnis oder eine Meinung münden. Mal begeben wir uns bewusst in Gedankengänge hinein, mal laufen sie ganz automatisch ab. Versuchen Sie doch einmal, jetzt in diesem Augenblick, aufzuhören, logisch verknüpfte Gedanken zu spinnen. Stellen Sie Ihre Gedanken ab. Denken Sie nicht darüber nach, was Sie gleich tun werden, wenn Sie dieses Kapitel fertig gelesen haben, oder warum

Ihr Freund Sie gestern Abend nicht angerufen hat. Na los, versuchen Sie es. Sie werden es nicht schaffen.

In Anbetracht der Tatsache, dass wir im wachen Zustand fast ununterbrochen denken, sollte uns interessieren, wie das funktioniert. Die Grundeinheit eines Gedankens ist eine Proposition – eine (logische) Aussage, die festgestellt oder bestritten werden kann. Auch eine bildliche Vorstellung oder ein Gefühl ist, genau genommen, eine Proposition. Beides hat eine Bedeutung und kann festgestellt oder bestritten werden und mit anderen Wahrnehmungen und Gefühlen logisch kombiniert und verknüpft werden. Wenn Sie sich beispielsweise bildlich vorstellen, wie Jack und Jill einen Berg hinaufsteigen, kann die inhaltliche Bedeutung (Proposition) dieser Vorstellung so ausgedrückt werden – „Jack und Jill steigen auf einen Berg." Und wenn ich eine andere Person besonders gern habe, kann ich die Bedeutung meines Gefühls so ausdrücken – „Ich liebe dich."

Eine Proposition ist zu unterscheiden von einem Satz, der die Proposition(en) transportiert. Der Satz „Auf den Berg hinauf steigen Jack und Jill" enthält dieselbe Proposition wie „Jack und Jill steigen auf einen Berg". Der Satz „Du wirst von mir geliebt" bedeutet inhaltlich genau dasselbe wie „Ich liebe dich".

Der Satz „Ich liebe dich" kann aber auch gebraucht werden, um viele verschiedene inhaltliche Bedeutungen (Propositionen) zu transportieren, je nachdem, wer mit den Pronomen „du" und „ich" gemeint ist. Er könnte beispielweise bedeuten „Denise liebt Robert" oder „Angelina liebt Brad", je nachdem, wer gerade „Ich liebe dich" denkt oder sagt. Der springende Punkt bei der Sache ist der, dass in

jedem dieser Fälle (Äußerungskontexte) die ausgedrückte Proposition entweder wahr ist oder falsch.

Wenn wir über etwas nachdenken, entwickeln wir eine Reihe von Propositionen, die logisch miteinander verbunden sind. Dasselbe tun wir, wenn wir eine andere Person von unserer Meinung zu überzeugen suchen. Es geht darum, dass das, was wir sagen, *logisch miteinander verbunden* sein muss. Propositionen, die nicht logisch miteinander verbunden sind, mögen amüsant sein, sind aber nicht überzeugend. Wir sprechen dann von einem „Wörtersalat" oder tun es abfällig als „völlig bescheuert" ab. Und wenn jemand darin versagt, sich in sinnhafter Weise verständlich zu machen (oder wir uns gar selbst nicht mehr verstehen), sind wir rasch besorgt und überlegen vielleicht, professionelle Hilfe zu holen und einen Psychiater zu Rate zu ziehen.

Propositionen müssen also logisch miteinander verbunden sein – aber was heißt das genau? Im Kern geht es in der Logik darum, wie die Wahrheit einiger Propositionen mit der Wahrheit anderer verbunden ist. Eine Reihe logisch miteinander verbundener Propositionen nennt man ein *Argument*. Oder präziser formuliert: Ein Argument ist eine Reihe von zwei oder mehr Propositionen, die so aufeinander bezogen sind, dass alle genannten Aussagen auf einen logischen Schluss hinführen (außer der letzten, versteht sich). Die vor dem Schluss genannten Aussagen sind dabei die Prämissen, die letztgenannte Aussage ist die Konklusion oder Schlussfolgerung. Hier ein einfaches Argument:

Prämisse: Jack und Jill stiegen auf einen Berg.
Konklusion: Daraus folgt, Jack stieg auf einen Berg.

In diesem einfachen Argument muss die Prämisse die Konklusion stützen. Es ist ein ziemlich gutes Argument. Die Konklusion folgt logisch aus den Prämissen, sie ist also deduktiv valid (gültig) – die Prämisse als wahr anzunehmen und die Konklusion als falsch abzulehnen, würde einen Widerspruch bilden und ist logisch ausgeschlossen. Wenn Sie die Prämisse für wahr halten, müssen Sie der logischen Notwendigkeit folgend auch den Schluss eines deduktiv validen Arguments für wahr halten.

Die Überführung der Prämissen hin zu einer Konklusion – dem logischen Zusammenhang zwischen beiden – wird als *Inferenz* (Folgerung) bezeichnet, auf der das Argument aufbaut. Man muss kein Logiker sein, um die inferenziellen Zusammenhänge zwischen den vorgebrachten Propositionen zu erkennen, wenn Sie Ihren Freund in seinem angeblich kaputten Auto auf den Parkplatz vor dem Kino fahren sehen. Aber welche Inferenzkette bräuchte es, um die folgende Geschichte zu verstehen?

> Mary verstaut die Picknicksachen im Kofferraum ihres Autos.
> Die Autofahrt dauert über eine Stunde.
> „Oh, nein", denkt sie auf einmal. „Das Bier wird eine warme Brühe sein."

Für jemanden, der überhaupt gar nicht weiß, was ein Picknick ist, wäre dies keine Geschichte. Es wären drei zusammenhanglose Sätze – reinster Wörtersalat eben. Aber Sie hatten offenbar kein Problem, diese Geschichte zu verstehen, weil Sie natürlich wissen, was ein Picknick ist. Und

so haben Sie in die Struktur der Geschichte weitere Propositionen gefüllt, die logisch zusammenhängen. Etwa so:

> Mary verstaut die Picknicksachen im Kofferraum ihres Autos.
> + Picknicks macht man im Sommer.
> + Im Sommer ist es heiß.
> + Zu einem Picknick gehört Bier.
> Die Autofahrt dauert über eine Stunde.
> + Sachen, die im Kofferraum liegen, werden warm.
> „Oh, nein", denkt sie auf einmal. „Das Bier wird eine warme Brühe sein."

Die Propositionen mit einem Plus-Zeichen davor, sind die, die Sie der Geschichte automatisch hinzugefügt haben, ausgehend von dem, was Sie über Picknicks wissen. Für jemanden, der weiß, was ein Picknick ist, besteht diese Geschichte aus einer Reihe logisch miteinander verbundener Propositionen. Sie bildet in der Tat ein deduktiv valides Argument. Alle Prämissen als wahr anzunehmen, die Konklusion aber als falsch abzulehnen (die letzte Zeile der Geschichte), würde einen Widerspruch bilden. Wenn ein Argument deduktiv valid ist, garantiert die Wahrheit der Prämissen die Wahrheit der Konklusion.

Für die deduktive Gültigkeit kommt es also darauf an, dass die Prämissen wahr sind. Oder ganz formal definiert: Sind alle Prämissen wahr, dann muss auch die Konklusion wahr sein. Betrachten wir folgendes Argument:

> Prämisse: Der Mond ist aus Schweizer Käse gemacht.
> Prämisse: Alle Schweizer Käse werden in Molkereien gemacht.
> Konklusion: Der Mond wurde in Molkereien gemacht.

Dieses Argument ist valid – das heißt, wenn beide Prämissen wahr sind, dann ist auch die Konklusion wahr. Um genauer zu sein: Wir würden sagen, dass dieses Argument deduktiv valid ist, aber nicht vernünftig. Ein *vernünftiges* Argument ist ein deduktiv valides Argument, basierend auf wahren Prämissen.

Wenn ein Argument bei Wahrheit der Prämissen die Wahrheit der Konklusion zwar nicht garantiert, aber dennoch *wahrscheinlich* bzw. *plausibel* macht, beinhaltet es eine sogenannte *induktive* Inferenz.

Ein induktives Argument gelingt, wann immer seine Prämissen einen legitimen Beweis oder eine Abstützung für die Wahrheit seiner Konklusion liefern. Es wäre jedoch nicht völlig inkonsistent, sich mit einem Urteil zurückzuhalten oder die Konklusion gar zu leugnen. Analoge Schlüsse beispielsweise stützen sich auf induktive Inferenz. Ein analoges Argument sieht so aus:

> Prämisse 1: Objekt X und Objekt Y ähneln sich in den Eigenschaften Q_1 bis Q_n.
> Prämisse 2: Objekt X weist zusätzlich Eigenschaft P auf.
> Konklusion: Objekt Y weist ebenfalls Eigenschaft P auf.

Beachte: Die Wahrheit der Prämissen garantiert nicht, dass die Konklusion ebenfalls wahr sein muss. Induktive Argumente werden nicht auf Basis der Validität bewertet, sondern stattdessen nach der Stärke der Beweise in den dargelegten Prämissen. Ein starkes induktives Argument ist eines, dessen Konklusion auf vielen schlüssigen Beweisen basiert. Ein schwaches induktives Argument basiert auf wenigen oder schwachen Beweisen.

Eine Reise in die Welt der Logik

Nun habe ich mit dem Ausdruck „logisch miteinander verknüpft" ziemlich um mich geworfen, ihn aber nicht wirklich erklärt. Um dies nachzuholen, müssen wir zunächst verstehen, was Logik ist.

Logik ist ein Zweig der Mathematik. In der Mathematik drücken wir mathematische Beziehungen aus, indem wir Zeichen verwenden, und diese Zeichen machen wir ineinander überführbar durch Regeln, die diese mathematischen Beziehungen aufrechterhalten. Hier ein Beispiel: Angenommen, Sie haben zwei Äpfel. Jemand gibt ihnen noch einen Apfel dazu. Wie viele haben Sie jetzt? Sie können die Äpfel einfach abzählen oder eine kleine Kopfrechnung anstellen. Oder sie verfahren formal und symbolisieren Ihre beiden Äpfel durch ein Zeichen: 2. Dass Sie einen weiteren Apfel hinzubekommen, symbolisieren Sie dann durch ein weiteres, recht praktisches Zeichen: +. Den Apfel, den Sie hinzubekommen, symbolisieren Sie wiederum mit einem Zeichen: 1. Da Sie wissen wollen, wie viele Äpfel Sie nun haben, drücken Sie das Ergebnis durch das Gleichheitssymbol aus: =. Und schließlich können Sie das tatsächliche Ergebnis durch ein weiteres Zeichen darstellen: x. Das ganze Ereignis stellt sich formal nun wie folgt dar: $2 + 1 = x$. Und schon sind wir im Land der Mathematik angekommen, wo nur Zeichen und Regeln existieren.

Um das mathematische Ergebnis unserer kleinen Rechnung zu finden, müssen wir sicherstellen, dass die Zeichen, die wir gewählt haben, auch die Bedeutung tragen, die wir beabsichtigen. In unserem Fall repräsentieren die Zeichen „1" und „2" ganze Zahlen auf dem Zahlenstrahl.

Sie repräsentieren zudem Mengen. Zudem brauchen wir einige Regeln, damit wir diese Zeichen so kombinieren, dass sie uns die richtige Lösung liefern. Wir brauchen also eine Regel für „+“, für das Plus-Zeichen, und die basiert auf dem Vorgang des Zählens oder der Additionsfunktion. Eine mathematische Funktion ist eine Beziehung zwischen zwei Mengen, in der die Elemente der einen Menge in einer Eins-zu-eins-Beziehung oder einer Eins-zu-viele-Beziehungen den Elementen der anderen Menge zugeordnet sind. Die Additionsfunktion nimmt zwei oder mehr Zeichen aus einer Eintragsmenge (Definitionsmenge) und ordnet sie genau einem Zeichen in der Ergebnismenge (Zielmenge) zu. In unserem Falle ordnet sie „2 + 1“ dem Zeichen „3“ zu.

War's das? Natürlich nicht. Wir werden die Welt der Mathematik nun verlassen, indem wir unser Ergebniszeichen auf ein Beispiel in der realen Welt übertragen und es interpretieren. Eben haben wir von Äpfeln gesprochen. Also interpretieren wir das Zeichen „3“ als Verweis auf die Menge – auf drei Äpfel also. Und das war's!

Es gilt in diesem Spiel also, drei Schritte zu vollziehen:

1. Man übertrage ein Ereignis aus der realen Welt in Zeichen aus der mathematischen Welt.
2. Man wende die entsprechende Regel aus der mathematischen Welt auf diese Zeichen an, um eine Lösung zu erhalten.
3. Man nehme die Lösung und übertrage sie zurück in die reale Welt.

„Ganz schön viel Arbeit, nur um herauszufinden, dass zwei Äpfel plus ein Apfel drei Äpfel sind“, höre ich Sie sagen.

Und Sie haben Recht. Aber was, wenn Sie 2 378 425 Äpfel haben und ich Ihnen 45 823 Äpfel dazugebe? Mit (Ab) Zählen kommen Sie dann nicht mehr weit. Aber wenn Sie wissen, wie Sie dieses Problem in die Welt der Mathematik übertragen können, ist die Aufgabe leicht – so leicht, dass sie automatisiert werden kann. Jeder Taschenrechner liefert Ihnen das Ergebnis blitzschnell.

Sie müssen nur folgende gedankliche Kurve kriegen: Wir versuchen jetzt mit Sätzen das zu tun, was die Mathematik mit Zahlen tut – *wir formalisieren die Sätze als Zeichen, die sie durch Regeln ineinander überführbar macht*.

Die einfachste Form der Logik, die wir hier diskutieren, ist die Aussagenlogik. Ebenso wie mathematische Zeichen, das Plus-Zeichen (+) etwa oder das Gleichheitszeichen (=), für die mathematischen Funktionen der *Addition* und der *Gleichheit* stehen, stehen für die Wahrheitsfunktionen in der Aussagelogik Zeichen wie ¬ für die logischen Funktionen der *Negation* und ⩒ für ein sogenanntes *exklusives Oder* (ein logisches, ausschließliches Oder). Die arithmetische Funktion des Zusammenzählens (des Plus-Rechnens) erfordert das Rechenzeichen +, das vor den einzelnen Ziffern der Eintragsmenge steht, woraus sich eine Zahl ergibt, die die Summe der Zahlen aus der Eintragsmengen repräsentiert – wie in unserem Beispiel mit den Äpfeln.

In der Aussagenlogik gibt es Variablen (z. B. P und Q), die für Elementarsätze (also elementare, nicht weiter zerlegbare Aussagen) stehen. Solche Elementarsätze werden durch bestimmte Symbole (sogenannte logische Operatoren) zu Formeln verknüpft, die eine Proposition ausdrücken. Funktionen (sogenannte Wahrheitsfunktionen) ordnen diesen Propositionen dann Wahrheitswerte zu. Der

Wahrheitswert einer Proposition kann nur die beiden Zustände „wahr“ oder „falsch“ annehmen.

Im obigen Beispiel mit dem kaputten Auto und der Mitfahrgelegenheit zum Kino, lief Ihre Gedankenkette so ab:

> Er hat es entweder prompt repariert oder er hat mich angelogen.
> Aber zum Reparieren blieb ja kaum Zeit.
> Also hat er mich angelogen.

Ist dies ein valides Argument? Wenn Sie Ihren Freund mit diesen Gedanken konfrontieren würden, hätten Sie ihn damit hieb- und stichfest überführt oder könnte er Ihnen vorwerfen, unlogisch zu sein? Wir können Wahrheitsfunktionen verwenden, indem wir in denselben drei Schritten vorgehen wie im Beispiel mit den Äpfeln. In diesem mathematischen Problem haben wir die Menge der Äpfel in Zeichen übertragen. Hier übertragen wir nun Propositionen in Zeichen. So wie wir das Zeichen 2 für „zwei Äpfel“ verwendet haben, soll nun P für „prompt repariert“ stehen und Q für „hat mich angelogen“. Das $\dot{\vee}$ -Zeichen repräsentiert „Oder“, das ¬-Zeichen ein „nicht“ (Negation). Damit haben wir die Zeichen für die Wahrheitsfunktionen in diesem Argument festgelegt:

$$\begin{array}{l} P \mathbin{\dot{\vee}} Q \\ \underline{\neg P} \\ \hline Q \end{array}$$

Sobald wir die Elementarsätze des Arguments in Zeichen aus der mathematischen Welt übertragen haben, können die numerischen Zeichen 2 und 1 für alles Mögliche stehen – zwei Äpfel, zwei Computer, zwei Weltkriege usw. Das + steht für eine und nur für eine Sache: die Additionsfunktion, die die Summe ihrer Eintragsmengen abbildet. Das „Äpfel-Problem" haben wir mithilfe der Additionsfunktion gelöst. In der Welt der Aussagenlogik mit ihren Wahrheitsfunktionen können P und Q für jedwede einfache Proposition stehen – „Er hat das Auto prompt repariert", „Fido ist ein Hund", „Der Mond besteht aus Schimmelkäse", „Ich mag Schokolade" usw.

Nun wenden wir die Wahrheitsfunktion für „oder" (das logische, ausschließliche Oder) an, was folgendermaßen aussieht (wobei T für „wahr" steht, F für „falsch"):

P	Q	P $\dot{\vee}$ Q
T	T	F
T	F	T
F	T	T
F	F	F

Diese Wahrheitsfunktion besagt, dass eine Aussage dieser Form nur dann wahr ist, wenn eine Proposition wahr ist und die andere falsch. Dies entspricht dem, was wir gemeinhin unter „oder" in einer normalen Unterhaltung verstehen. Wenn die Bedienung im Restaurant Sie fragt, ob Sie lieber Suppe *oder* Salat möchten, dann meint sie, dass Sie nicht beides haben können. Sie müssen sich für das eine *oder* das andere entscheiden.

Nun wenden wir die Wahrheitsfunktion für die Negation an. Diese Wahrheitsfunktion ist leicht; es nimmt einfach eine Proposition an und gibt ihr Gegenteil an. Das sieht folgendermaßen aus:

P	¬P
T	F
F	T

Schauen wir uns nun an, ob dieses Argument valid ist, wenn wir diese Wahrheitsfunktionen verwenden:

Elementarsätze		Prämissen		Konklusion
P	Q	P $\dot{\vee}$ Q	¬P	Q
T	T	F	F	T
T	F	T	F	F
F	T	T	T	T
F	F	F	T	F

Der Wahrheitswert eines Arguments ist valid, wenn in jeder Tabellenzeile, in der die Prämissen wahr sind, auch die Konklusion wahr ist.

Wenn also die Aussage „P oder Q“ wahr ist und die Aussage „P ist falsch“ ebenfalls wahr ist (dritte Zeile der Tabelle), dann muss Q wahr sein. Das Argument ist damit valid.

Die Aussagenlogik verwendet logische Operatoren, die sich in unserer natürlichen Alltagssprache wiederfinden, wie: *und*, *oder*, *wenn*, *nur wenn*, *außer wenn, nicht*. Diese logischen Grundfunktionen werden verwendet, um die Validität (Gültigkeit) eines Arguments zu überprüfen, so wie in unserem Beispiel mit dem exklusiven Oder.

In der *Prädikatenlogik ersterStufe* (auch als *Quantorenlogik* bezeichnet) können wir sogenannte *Quantoren* verwenden, um Propositionen zu erhellen (Quantifizierung) und damit ihre *Prädikate* (oder *Funktoren*) auszudrücken, die das Element, auf das sie sich beziehen, näher bestimmen. Zum Beispiel können wir der Proposition „John hat sein Auto repariert" mit dem Prädikat F(j,c) Ausdruck verleihen, wobei F für „repariert" (*fixed*) steht, j für „John" und c für „Auto" (*car*). Wir können auch Zeichen verwenden, die Mengen- oder Ereignisreihen repräsentieren, zum Beispiel „Jemand hat das Auto repariert". Dies sieht dann so aus:

$$\exists(x)F(x,c)$$

Das Zeichen ∃ ist der sogenannte Existenzquantor; er zeigt an, dass eine Menge existiert, die mindestens ein Element in sich trägt. Wir können diesen Ausdruck demnach wie folgt lesen: „Es gibt mindestens ein x (eine Sache oder eine Person) so, dass dieses x das Auto repariert hat." Wir können auch etwas ausdrücken wie „Jeder repariert das Auto":

$$\forall(x)F(x,c)$$

Das Zeichen ∀ ist der universale Allquantor, gelesen als „für jedes x (gilt)", was in vielen Fällen dasselbe ist wie „für alle x (gilt)". Wollen wir also den Gedanken ausdrücken „Alle Hunde sind Tiere", tun wir dies folgendermaßen – wobei D für „Hunde" (dogs) und A für „Tiere" (animals) steht:

$$\forall(x)(Dx \rightarrow Ax)$$

In Worten heißt dies: „Nimm, was immer du möchtest – solange es ein Hund ist, ist es ein Tier."

Einmal mehr gilt, dass alle Bedeutung (oder aller Inhalt) verloren geht, sobald wir uns in der Welt der Logik befinden. Ebenso wie die mathematische +-Funktion ist die $\dot{\vee}$-Funktion „blind" für alles, außer der Form der Zeichen, die sie auf andere Zeichen abbildet. Nachdem wir die Zeichen in Funktionen überführt haben, müssen wir diese wieder zurück in ihre Bedeutung in der realen Welt übertragen.

Die Prädikatenlogiken höherer Stufe sind dazu da, um Ideen zu erfassen, die durch Aussagenlogik oder *Prädikatenlogik erster Stufe* nicht ausgedrückt werden können. Diese Logiken nämlich können zum Beispiel Aussagen wie „es ist notwendig, dass" oder „es ist möglich, dass" nicht ausdrücken. Doch solche Wendungen gebrauchen wir im Alltag ständig und wir verwenden sie auch in Argumenten. Der Zweig der Logik, der sich mit den Folgerungen um die Modalbegriffe „notwendig" und „möglich" befasst, heißt *Modallogik*. Mithilfe der Modallogik können wir Notwendigkeiten ausdrücken, indem wir das Zeichen $\Box$ verwenden, und Möglichkeiten, indem wir das Zeichen $\Diamond$ verwenden. Die Validität von Argumenten, die diese Ideen enthalten, können nicht mittels Wahrheitstabellen überprüft werden, denn Wahrheitstabellen erfassen nur Wahrheitsfunktionen. Stattdessen überprüft man Modalargumente mittels einer Formalisierung des Begriffs der *möglichen Welt*; man konstruiert eine Reihe möglicher Welten und fragt dann, ob eine bestimmte Proposition (Aussage) für alle oder eine von ihnen wahr ist. Wenn sie für alle wahr ist, ist diese Proposition *notwendig wahr*. Wenn sie nur für einige wahr ist,

ist sie *möglich wahr*. Wenn sie für alle nicht wahr ist, ist sie notwendigerweise falsch. Betrachten wir zum Beispiel die Aussage „Alle Kreise sind rund". Diese Aussage wird sich für alle möglichen Welten, die Sie konstruieren wollen, als wahr erweisen. Die Aussage „Alle Kreise sind rot" jedoch kann für einige Welten wahr, für andere falsch sein. Folglich ist die Aussage „Alle Kreise sind rund" notwendig wahr für jedes Argument, das Sie konstruieren, wohingegen „Alle Kreise sind rot" nur möglich wahr ist. Die Aussage „Alle Kreise sind Quadrate" ist für alle möglichen Welten nicht wahr und folglich notwendig falsch für jedes Argument, das Sie konstruieren. Nun verfahren Sie mit den möglichen Welten so, wie wir das mit unseren Wahrheitstabellen gemacht haben: Sind alle Prämissen für jedes mögliche Modell wahr, muss auch die Konklusion wahr sein, und das Argument ist damit valid (gültig).

Und schließlich gibt es noch Bedeutungen, die spezielle Modallogiken erfordern. Diese finden sich meist in Bereichen, die sehr theorielastig sind. Betrachten wir zum Beispiel folgendes Argument:

> Wenn die Bremse betätigt wird, wird das Auto langsamer.
> Das Bremspedal wird getreten.
> Daraus folgt, das Auto wird langsamer.

Wollten wir dieses Argument als Wahrheitsfunktion ausdrücken, wäre es valid. Dieser Argumenttyp wird als Modus ponens bezeichnet. Er ist aufgrund der Wahrheitsfunktion für bedingende Aussagen der Form „wenn-dann" immer valid.

Elementarsätze		Prämissen		Konklusion
P	**Q**	**P → Q**	**P**	**Q**
T	T	T	T	T
T	F	F	T	F
F	T	T	F	T
F	F	T	F	F

Wie Sie an der Wahrheitstabelle sehen können, gilt, wenn beide Prämissen wahr sind, muss auch die Konklusion wahr sein. Aber ergibt das aus kausaler Sicht einen Sinn? Was, wenn die Bremsleitungen durchgeschnitten oder die Straßen eisglatt sind? Oder die Bremsscheiben verschlissen? Es gibt viele Faktoren, die in Betracht gezogen werden müssen, bevor diese Schlussfolgerung gezogen werden kann. Des Weiteren ist der Wenn-dann-Operator in diesem Argument nicht einfach ein bedingender logischer Operator mit einem Wahrheitswert Er beschreibt eine kausale Beziehung und Kausalzusammenhänge können nicht durch eine Wahrheitsfunktion erfasst werden. Wenn wir also kausal denken, tun wir etwas sehr viel komplexeres, dass mit der Aussagenlogik nicht adäquat erfasst werden kann. Wir konstruieren eher so etwas wie mögliche Welten mit einem speziellen Dreh.

Die *Kausallogik* ist ein Zweig der Modallogik, der kausale Beziehungen betrachtet (mehr dazu in Kapitel 6). Kausale Beziehungen, auch als Ursache-Wirkungs-Beziehungen bezeichnet, werden in Form von notwendigen Bedingungen ausgedrückt (ein bestimmter kausaler Faktor muss gegeben sein, damit eine bestimmte Wirkung unter den gegebenen Umständen zustande kommt), sowie in Form von hinrei-

chenden Bedingungen (wenn ein bestimmter kausaler Faktor gegeben ist, dann folgt eine bestimmte Wirkung garantiert). In unserem Beispiel ist das Treten des Bremspedals weder eine notwendige noch eine hinreichende Bedingung, damit das Auto langsamer wird. Betrachten wir das folgende Argument:

> Wenn Sie einen gültigen Bibliotheksausweis besitzen, dann können Sie sich ein Buch aus der Bibliothek ausleihen.
> Sie haben keinen gültigen Bibliotheksausweis.
> Daraus folgt, Sie können sich kein Buch ausleihen.

In der Aussagenlogik ist dies ein nicht valides (ungültiges) Argument. (Das können Sie mir ruhig glauben oder Sie rechnen anhand der Wahrheitstabelle nach, um sich selbst zu überzeugen). Aber zugegeben, es sieht nach einem stichhaltigen Argument aus. Das liegt daran, dass es in diesem Argument um Erlaubnis und Verpflichtung geht. Und das ist der Bereich der *deontischen Logik.* Sie ist ein Zweig der Modallogik und studiert die logischen Relationen zwischen „erlaubt", „geboten" und „verboten", die durch deontische Modalitäten charakterisiert sind. (Der Begriff „deontisch" leitet sich ab vom altgriechischen Wort *déon*, was so viel heißt wie das Nötige, das Angemessene.) Diese Logik führt zwei neue formale Zeichen ein: OA steht für eine Verpflichtung (A zu tun) und PA steht für die Erlaubnis (A zu tun). Sie sind mit Bezug aufeinander definiert: Wenn du verpflichtet bist, etwas zu tun, ist es dir nicht erlaubt, es nicht zu tun. Und wenn es dir erlaubt ist, etwas zu tun, dann bist du nicht verpflichtet, es nicht zu tun.

Logiker erfinden neue Arten der Logik, wenn die bestehenden Logiken nicht ausreichen, legitime Interferenzen zu erfassen. Je ausgefeilter die Logiken werden, umso komplexere Argumente können bewertet werden. Für einfache Argumente wie hier in unseren Beispielen mag es uns scheinen, dass all diese großen Mühen am Ende doch nur herzlich wenig fruchten. Aber so erschien es uns zunächst auch in unserem Beispiel „zwei Äpfel plus ein Apfel", woraufhin wir eines Besseren belehrt wurden, als wir 2 378 425 Äpfel plus 45 823 Äpfel abzählen wollten – da haben wir schnell begriffen, wie notwendig formalisierte Mathematik auf Zeichenebene ist. Dasselbe gilt für die logische Beweisführung in Argumenten. Gewiss, es kann schwierig sein, in sehr langen und komplexen Argumenten jeden einzelnen Schritt bis zur Schlussfolgerung nachzuverfolgen. Doch es kann die ganze Sache auch sehr viel einfacher machen, das Argument in die Welt der Zeichenlogik zu übertragen und die Regeln der Logik darauf anzuwenden.

Wie logisch denken wir wirklich?

Wie gut sind wir wirklich im deduktiven, sprich im schlussfolgernden Denken? Ich möchte Ihnen eine Studie vorstellen, die sich genau mit dieser Frage befasst und zu den typischen Ergebnissen gelangt (Evans et al. 1983). Die Teilnehmer bekamen die folgenden Instruktionen:

> In diesem Experiment soll die logische Denkfähigkeit getestet werden. Sie bekommen acht Aufgaben. Auf jeder Seite finden Sie zwei Aussagen. Entscheiden Sie bitte auf-

grund dieser Aussagen, ob daraus bestimmte Schlüsse (die unter den Aussagen stehen) logisch abgeleitet werden können. Treffen Sie Ihre Antworten in der Annahme, dass beide Aussagen tatsächlich wahr sind.
Sind Sie der Meinung, dass ein Schluss notwendigerweise folgt, kreuzen Sie „Ja" an, andernfalls „Nein". Lassen Sie sich Zeit, bis Sie sicher sind, die richtige Antwort gefunden zu haben.

Es folgen einige Beispiele von Syllogismen, die es zu befinden galt. Aber denken Sie daran: Treffen Sie Ihre Antworten in der Annahme, dass beide Aussagen (Prämissen) tatsächlich wahr sind und überlegen Sie, ob die Konklusion folglich ebenfalls wahr sein muss oder nicht.

Kein Polizeihund ist bissig.
Einige abgerichtete Hunde sind bissig.
Daraus folgt: Einige abgerichtete Hunde sind keine Polizeihunde.

Kein Nahrungsstoff ist teuer.
Einige Vitamintabletten sind teuer.
Daraus folgt: Einige Vitamintabletten sind keine Nahrungsstoffe.

Kein Suchtmittel ist teuer.
Einige Zigaretten sind teuer.
Daraus folgt: Einige Suchtmittel sind keine Zigaretten.

Kein Millionär arbeitet schwer.
Einige reiche Leute arbeiten schwer.
Daraus folgt: Einige Millionäre sind keine reichen Leute.

Und die Antwort lautet: Die ersten beiden Schlüsse sind valid, die beiden letzten sind nicht valid.Wenn Sie diese Antworten überraschend finden, dann liegt dies wahrscheinlich daran, dass Sie der sogenannten *Antwortverzerrung* (oder Antwortfehler) aufsitzen: Sie geben der ersten Information ein stärkeres Gewicht als der (konsequenten) logischen Form bei der Beurteilung der Syllogismen. Die Konklusionen des ersten und des dritten Syllogismus sind glaubhaft. Die Konklusionen des zweiten und des vierten Syllogismus sind nicht glaubhaft.

Die Probanden dieser Studie zeigten eine starke Tendenz zur Antwortverzerrung. Wären die Syllogismen einzig aufgrund der logischen Form (der logischen Verknüpfung von Prämissen und Konklusion) zu beurteilen gewesen, hätte sich für die validen Syllogismen eine Akzeptanzrate von 100 % und für die nicht validen Syllogismen eine Akzeptanzrate von 0 % ergeben. Aber das ist nicht passiert. Stattdessen schienen die Probanden beides in ihre Überlegungen miteinzubeziehen – Glaubhaftigkeit und logische Form –, was sie häufig zu falschen Urteilen führte. Valide Syllogismen mit glaubhaften Konklusionen wurden im gesamten Studienverlauf zu rund 85 % korrekt angenommen, während diejenigen mit nicht glaubhaften Konklusionen nur zu rund 55 % korrekt angenommen wurden. Die Akkuranz ließ nach, wenn Validität und Aufgabenurteil in einander widersprüchliche Entscheidungen mündeten! In gleicher Weise wurden nicht valide Syllogismen mit nicht glaubhaften Konklusionen im gesamten Studienverlauf zu lediglich rund 10 % korrekt angenommen, diejenigen mit glaubhaften Konklusionen hingegen zu rund 70 %! Auch hier zeigte sich wieder eine signifikante Reduktion der Akkuranz.

Die Teilnehmer zeigten eine größere Tendenz zu Antwortverzerrungen, wenn sie ihre Antworten unter Zeitdruck treffen mussten. Evans und Curtis-Holmes (2005) stellten den Teilnehmern ihrer Studie frei, sich mit der Beurteilung der Syllogismen beliebig viel Zeit zu lassen oder ihre Antwortzeit auf zehn Sekunden zu begrenzen. War die Antwortzeit begrenzt, wurden valide Syllogismen mit nicht glaubhaften Konklusionen im gesamten Studienverlauf zu knapp 40 % korrekt angenommen, invalide Syllogismen mit nicht glaubhaften Konklusionen hingegen zu fast 80 %! Unter Zeitdruck waren die Probanden eher geneigt, sich in der Beurteilung der Argumente einfach auf die erstgenannten Informationen zu verlassen, anstatt die logische Verknüpfung der Aussagen zu analysieren.

Im vorangegangen Kapitel haben wir gesehen, dass Prozesse der Urteilsbildung in zwei neurologisch getrennten Systemen ablaufen, in System 1 und System 2, wie sie in der Kognitionswissenschaft bezeichnet werden. Diese Zwei-Prozess-Theorie, wie sie dort heißt, wird herangezogen, um eine Vielzahl kognitiver Phänomene zu erklären. System 1 liefert schnelle, automatische, (meist) emotional gesteuerte Urteile und läuft außerhalb der bewussten Wahrnehmung ab. Wenn Sie ein Bauchgefühl haben oder Ihnen irgendetwas urplötzlich, aus unerklärlichen Gründen, in den Sinn schießt, dann war System 1 aktiv. Im Unterschied dazu arbeitet System 2 langsamer und kontrollierter und liefert Urteile, die auf bewussten, mentalen Aktivitäten gründen. Wenn die Ergebnisse dieser beiden Systeme übereinstimmen, kann ein Urteil relativ rasch und mit einem hohen Maß an Sicherheit gebildet werden. Kommen die beiden Systeme aber zu unterschiedlichen Ergebnissen, muss der

Konflikt gelöst werden, was den Urteilsprozess erheblich verlangsamt. Die Antwortverzerrung zeigt, dass die Ergebnisse von System 1 nicht selten über die Ergebnisse von System 2 dominieren.

Das zweigeteilte, theoretische Systemmodell zur Strukturierung der Denkprozesse wird umfassend gestützt durch Erkenntnisse der neurowissenschaftlichen Forschung. Im Rahmen etlicher Studien nutzen die Forscher Methoden der medizinischen Diagnostik zur Messung der elektrischen Aktivität im Gehirn – wie beispielsweise der schematischen Darstellung des Verlaufs *ereigniskorrelierter Potenziale* (ERP) mittels EEG (Elektroenzephalogramm) oder fMRT (funktionelle Magnetresonanztomografie) –, während die Probanden aufgefordert waren, Syllogismen und Argumente zu beurteilen (Goel & Dolan 2003; Luo et al. 2008). Dabei zeigten sich verschiedene Areale in den Frontallappen aktiv, ganz gleich, ob die Probanden korrekte oder nicht korrekte Urteile abgaben. Eine sehr hohe Aktivität zeigte sich im anterioren cingulären Cortex (ACC), wenn die Probanden einen Syllogismus oder ein Argument zu verarbeiten hatten, dessen logische Form dem eigenen Urteil zuwiderlief, und die registrierenden Systeme daher miteinander in Konflikt gerieten. Wenn die Probanden der Antwortverzerrung anheimfielen, zeigte sich der ventromediale präfrontale Cortex (VMPC) aktiv, während der dorsolateraler Präfrontalcortex (dlPFC) aktiv war, wenn sie logikbasiert urteilten.

Doch wie schaffen wir es, logischer zu denken? Es hat sich gezeigt, dass Disziplinen, die einen intensiven Gebrauch von Zeichen und Zeichenverknüpfungen erfordern, die menschliche Fähigkeit schulen, logische Formen zu erkennen und zu beurteilen. In einer Studie (2007) legten

Inglis und Simpson ihren Probanden einfache Argumente nach der *Prädikatenlogik erster Stufe* vor, die in ihrer Glaubhaftigkeit differierten, und baten sie, deren Validität zu beurteilen. Es gab zwei Gruppen von Probanden. Die eine bestand aus Studienanfängern, die ihr Studium der Mathematik an einer britischen Eliteuniversität gerade erst aufgenommen hatten, die andere aus Studienreferendaren, die sich nicht auf Mathematik spezialisiert hatten. Die Ergebnisse zeigten, dass die Mathematikstudenten sechsmal weniger der Antwortverzerrung anheimfielen als die Studienreferendare!

Was tun, wenn sich unsere Welt (sprich, unser Verstand) verändert?

Im realen Leben stoßen wir sehr viel häufiger auf Argumente, die extrem überzeugend aber nicht deduktiv valid sind. Wir haben diese Argumente oben als induktive Argumente beschrieben. Nun wollen wir einen Schritt weiter gehen. Sehr häufig sind stützende Beziehungen zwischen Prämissen und Konklusion nicht definitiv und können durch additionale Informationen potenziell bezwungen werden. Unter diesen Umständen muss ein intelligenter Denker jederzeit darauf vorbereitet sein, Konklusionen angesichts widersprüchlicher Informationen zurückzunehmen. Pollock beschreibt diesen Vorgang als *defeasible reasoning* (1987): „In einem Argument werden Prämissen als Gründe angeführt, um seine Konklusion zu rechtfertigen. Ein Argument ist insofern *defeasible*, als seine Konklusion durch eine Er-

weiterung der Prämissenmenge, also durch Hinzufügung zusätzlicher (Gegen)Gründe, unbegründet werden kann."[1]

Im Bereich der künstlichen Intelligenz spricht man auch von einem monotonen beziehungsweise nichtmonotonen Schließen, um diese Unterscheidung zu treffen (McCarthy 1980). Beim monotonen Schließen werden Inferenzen (Folgerungen) aufgrund neuer Informationseinträge, aktueller Meinungen und Regelhaftigkeiten gezogen. Diese Inferenzen kommen wahren Urteilen oder gültigen Schlüssen gleich, die in die Wissensbasis eingespeist werden und dort verbleiben – Wissen „wächst" also in monotoner Weise an. Beim nichtmonotonen Schließen können gültige Schlüsse durch Hinzunahme weiterer Informationen und Prämissen zu ungültigen Schlüssen werden und werden aus der Wissensbasis entfernt. Dadurch wächst und schrumpft die Wissensbasis in dynamischer Weise.

Betrachten wir einmal folgende Situation, die Pollock 1987 beschrieb. Sie sehen etwas, das ein rotes Tuch zu sein scheint. Also schließen sie Folgendes:

Das Tuch sieht rot aus.
Daraus folgt: Das Tuch ist rot.

Doch dann stellen Sie fest, dass das Tuch durch rotes Licht angestrahlt wird. Also modifizieren Sie Ihre Konklusion:

Das Tuch wird durch ein rotes Licht angestrahlt.
Daraus folgt: Das Tuch kann rot sein, muss aber nicht.

[1] Wang PH (2003) Defeasibility in der juristischen Begründung. Nomos, Baden-Baden, S. 11

Pollock verwendet den Begriff *defeaters*, um sich auf Informationen zu beziehen, die eine Schlussfolgerung gänzlich widerlegen oder die folgernde Verknüpfung zwischen Prämissen und Konklusion unterminieren. Sowohl vorgängige Überzeugungen als auch neue Informationen können als *defeaters* fungieren. Wenn wir zulassen, dass unsere vorgängigen Meinungen über logische Formen dominieren, lassen wir uns auf den Vorgang des *defeasible reasoning*, wie oben beschrieben, ein.

Betrachten wir an dieser Stelle noch einmal das kausale Argument, das offenbar dem Argumenttyp *Modus ponens* zuzurechnen war:

> Wenn die Bremse betätigt wird, wird das Auto langsamer.
> Das Bremspedal wird getreten.
> Daraus folgt: Das Auto wird langsamer.

Cummins et al. (1991, 1995, 1997) konnten zeigen, dass Probanden weit weniger geneigt waren, Konklusionen dieses Argumenttyps anzunehmen als Konklusionen wie beispielsweise diese:

> Wenn sie das Glas mit bloßen Fingerspitzen berührt, werden ihre Fingerabdrücke auf dem Glas sein.
> Sie berührt das Glas mit bloßen Fingerspitzen.
> Daraus folgt: Ihre Fingerabdrücke sind auf dem Glas.

Warum? Weil wir uns jede Menge *defeaters* für das Bremsszenario vorstellen können, aber nicht für das Fingerabdruckszenario. Es gibt viele (alterative) Gründe, warum ein Auto langsamer werden kann, außer durch das Treten des Bremspedals: Der Sprit geht aus, es fährt bergauf, das Ge-

lände ist holprig usw. Es gibt ebenso viele Gründe, warum ein Auto nicht langsamer wird, obwohl das Bremspedal gedrückt wurde (Cummins bezeichnet sie als *disablers*). Die Bremsleitungen könnten durchgeschnitten sein, die Bremsflüssigkeit könnte ausgelaufen sein, die Straßen könnten eisglatt sein usw. Dies ist ein Argument, das sehr leicht widerlegt und damit unbegründet gemacht werden kann, denn es lässt viele alternative Gründe und viele sogenannte *disablers* zu.

Aber betrachten wir noch einmal das Fingerabdruckargument. Hier ist es sehr schwierig, einen alternativen Grund zu finden, denn einmal auf dem Glas, ist es kaum vorstellbar, dass diese oder jene Person das Glas nicht angefasst hat. Um Pollocks Terminologie anzuwenden, könnten wir also sagen, dass das erste Argument überzeugend, aber widerlegbar ist, das zweite einfach nur deduktiv valid.

All das führt uns sehr deutlich vor Augen, dass wir in unseren Schlussfolgerungen offenbar immer versuchen, in unserer Wissensbasis Wahrheit und Konsistenz aufrechtzuerhalten. Sich Fakten aus dem Gedächtnisspeicher zu ziehen, ist leichter und schneller getan als logische Formen zu extrahieren und zu evaluieren – wie wir beinahe täglich in TV-Quizshows wie *Jeopardy* sehen können, wo den Teilnehmern Antworten aus verschiedenen Kategorien präsentiert werden und sie möglichst schnell eine passende Frage auf eine vorgegebene Antwort formulieren sollen. Wenn mühsam gewonnene Urteile einer logischen Form zuwiderlaufen, obsiegen oftmals erstere, insbesondere wenn wir unter Zeitdruck reagieren müssen. Außerdem können vorgängige Überzeugungen sehr häufig ein vermeintlich deduktiv valides Argument unterminieren.

Wie wir in den nachfolgenden Kapiteln noch sehen werden, ist es Fluch und Segen zugleich, dass wir uns auf vorgängige Überzeugungen verlassen. Unsere Überzeugungen können uns von schlechten Entscheidungen abbringen, uns aber auch stur und dickköpfig machen.

Box 5.1 Wie Aristoteles dachte

Ein kategorischer Syllogismus ist ein Argument, das aus genau drei kategorischen Propositionen besteht (zwei Prämissen und eine Konklusion) und genau drei kategorische Aussagen verwendet, wovon jede genau zweimal verwendet wird.

Alle Gänse sind Vögel.
Alle Vögel haben Federn.
Daraus folgt: alle Gänse haben Federn.

Kategorische Syllogismen wurden einst von Aristoteles eingeführt und bildeten den Grundstein seines logisch folgernden Denksystems. Logiker des Mittelalters haben eine recht einfache Methode entwickelt, mit der sich die verschiedenen Formen, in denen ein kategorischer Syllogismus auftreten kann, bezeichnen lassen. Das oben angeführte Argument hat die Form AAA-1 und es ist deduktiv valid. Formal sieht dies für die einzelnen Aussagen so aus:

Alle A sind B.
Alle B sind C.
Daraus folgt: Alle A sind C.

Egal, wofür A, B oder C stehen, ein Argument dieser Form ist deduktiv valid. Hier ein weiteres Beispiel:

Alle P sind nicht M.
Einige S sind nicht M.
Daraus folgt: Einige S sind nicht P.

Dies ist ein nicht valides Argument der Form EOO-2.
Zum Beispiel:

Alle Hunde sind keine Katzen.
Einige Vögel sind keine Katzen.
Daraus folgt: Einige Vögel sind keine Hunde.

Natürlich sind einige Vögel keine Hunde. Tatsächlich sind alle Vögel keine Hunde. Doch zu diesem Schluss können Sie über die behaupteten Prämissen nicht gelangen. Die Prämissen verknüpfen nämlich nicht Vögel mit Hunden in irgendeiner logischen Weise. Sie besagen lediglich, dass es Hunde und Katzen verschiedene Tierarten sind und dass einige Vögel sich von Katzen unterscheiden. Das bedeutet nicht, dass sie sich von Hunden unterscheiden. Folglich garantieren die Prämissen nicht die Wahrheit der Konklusion.

Dies ist kein Lehrbuch über Logik und deshalb müssen sie die kategorischen Syllogismen der aristotelischen Logik auch gar nicht vollständig verstehen. Was Sie aber mitnehmen sollten, ist, *dass logische Validität ausschließlich von der Form des Arguments abhängt*. Der Inhalt der Propositionen ist irrelevant.

Die aristotelische Logik stand als eine Methode, Inferenzen zu erfassen, 2000 Jahre lang allein. Der deutsche Philosoph Karl von Prantl (1820–1888), der sich mit der Geschichte der Logik im Abendland befasste, geht sogar so weit zu behaupten, dass sämtliche neue Systeme der Logik, die Logiker nach Aristoteles hervorbrachten, konfus, dumm oder verkehrt seien (*Stanford Encyclopedia of Philosophy*). Die Schwierigkeit bestand darin, dass die aristotelische Logik alles, was ausgedrückt, was argumentiert und bewiesen werden kann, stark begrenzt. Wirksamere Methoden waren nötig, um die expressive Macht der natürlichen Sprache und die natürlichen Inferenzen erfassen zu können, und dieser Herausforderung stellten sich viele Philosophen des 19. und 20. Jahrhunderts.

Mitte des 19. Jahrhunderts entwickelte der englische Mathematiker, Logiker und Philosoph George Boole (1815–1864) den ersten algebraischen Logikkalkül, mit dem er die aristotelische Logik erweiterte, indem er zuließ, dass ein Argument viele Prämissen und Klassen enthielt. Dieser algebraische Ansatz wurde später von Alfred Whitehead und Bertrand Russell zurückgewiesen zugunsten eines anderen Ansatzes, der von Gottlob Frege entworfen wurde und Gebrauch machte von logischen Konnektiven, Relationszeichen und Quantifikatoren. Russell und Whitehead hatten ein ehrgeiziges Ziel, und zwar suchten sie nachzuweisen, dass „die gesamte reine Mathematik aus rein logischen Prämissen folgt und sich alle in ihr auftretenden Grundbegriffe rein logisch definieren lassen" (Russell 1959, S. 74). Ihre Zusammenarbeit gipfelte in der Publikation des dreibändigen monumentalen Werks *Principia mathematica* (1910–1913) über die Grundlagen der Mathematik. Letztendlich aber erweis sich ihr ehrgeiziges Ziel als unerreichbar, als der österreichisch-amerikanische Mathematiker Kurt Friedrich Gödel zeigte, dass arithmetische Wahrheiten nicht aus einer Prämissenmenge abgeleitet werden können.

6
Was verursacht was?

Ein bisschen was von Kontrollfreak steckt in jedem von uns. Es drängt uns zu wissen, was wir anstoßen oder tun müssen, damit wir dieses oder jenes schöne Erlebnis oder Ereignis herbeiführen, oder auch dieses oder jenes Erlebnis vermeiden, auf das wir gut und gerne verzichten können. Wir möchten einfach wissen, was durch was verursacht wird, auch wenn die Dinge sind wie sie sind und wir daran nichts ändern können. Wir streben nach einem Zustand der kognitiven Erfüllung, der sich beispielsweise dann einstellt, wenn es uns wie Schuppen von den Augen fällt: „Aha! Darum ist das so." Um diese Erfüllung zu erlangen, vertrauen wir unwillkürlich auf ein sehr einfaches Grundkonzept, das der australische Philosoph John Mackie Mackie, J. (1974) den „Zement des Universums" nannte – die *Kausalität* (der Zusammenhang von Ursache und Wirkung).

Das Paradox der Kausalität

Aus psychologischer Sicht ist die Kausalität ein Paradox. Wir verwenden das Konzept der Kausalität, um die vielfältigen Ereignisse unseres täglichen Lebens zu hinterfragen

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_6,

und zu erhellen, wie etwa, warum das Auto heute nicht angesprungen ist oder manche Menschen gewalttätig sind. Die Ursachen jedoch entziehen sich unseren Sinnen – sie können nicht in unmittelbarer Weise wahrgenommen werden. Und genau das veranlasste den Philosophen David Hume (1711–1776) zu behaupten, Kausalität sei eine „Illusion" (Hume 1748). Wenn ein Ereignis (Ursache) ein anderes verursacht (Wirkung), so Hume, dann müssen diese beiden Ereignisse immer erstens zusammen auftreten (konstante Konjunktion), zweitens zeitlich eng aufeinanderfolgen, wobei die Ursache der Wirkung vorausgeht (temporale Priorität), drittens räumlich benachbart sein (räumliche Nähe), und schließlich muss viertens ein Ereignis die Kraft haben, das andere Ereignis hervorzubringen (notwendige Verknüpfung). Das Paradox dabei ist: Während wir die ersten drei Bedingungen in direkter Weise wahrnehmen können, können wir die vierte – die notwendige Verknüpfung zwischen den Ereignissen – nicht unmittelbar erkennen. Ein einfaches Beispiel: Stellen Sie sich vor, Sie beobachten, wie ein schwerer Hammer auf eine Kristallvase trifft mit der Folge, dass diese zerspringt. Lesen Sie diesen Satz jetzt noch einmal. Er enthält sämtliche Informationen über dasjenige Ereignis, das direkt wahrnehmbar ist. Sie müssten demnach so etwas im Kopf haben wie: „Der Hammer trifft die Vase, die dann zerspringt." Doch das denken Sie jetzt in diesem Moment nicht wirklich. Sie haben vielmehr die folgende Proposition im Kopf: „Der Hammer trifft die Vase und *verursacht*, dass sie zerspringt." Aber wie kommen Sie auf „verursacht"? Kausalität kann, wie gesagt, nicht in unmittelbarer Weise wahrgenommen werden. Eben darum behauptet Hume, dass Kausalität eine Illusion sein müsse, die

der menschliche Geist uns aufzwingt, in etwa so, wie unser visuelles System allerlei Illusionen unterliegt. Er schreibt: „Kraft und Notwendigkeit existieren im Geist, nicht in Gegenständen […] und sind daher Eigenschaften der Wahrnehmungen, nicht der Gegenstände, [sie] werden im Inneren durch die Seele empfunden, nicht äußerlich durch den Körper wahrgenommen“ (Hume 1748). Hume sah keine andere Möglichkeit, das Prinzip der Kausalität zu erklären, denn nicht zuletzt war er einer der Gründer des britischen Empirismus, einer philosophischen Strömung, die in ihrem Kern besagt, dass alles Wissen allein durch Sinneserfahrungen erlangt werden kann: Wir werden geboren als leeres, unbeschriebenes Blatt; alles, was wir wissen, haben wir durch sensorische Erfahrungen erlernt. Das Prinzip der Kausalität ist für eine solche Auffassung recht problematisch, denn ein kausaler Zusammenhang kann nicht direkt „erfahren“ werden. Er muss folglich eine Illusion sein, ein Trugbild – wie eine Fata Morgana in der Wüste.

Der deutsche Philosoph Immanuel Kant (1724–1804) wandte sich vehement gegen Hume und warf ihm vor, aus dem „*Causal*-Begriff“ einen „Bastard der Einbildungskraft, durch die Erfahrung beschwängert“ gemacht zu haben (Kant 1783). Er grenzt sich deutlich von ihm ab, indem er stattdessen argumentiert, dass die Existenz der Kausalität eine A-priori-Wahrheit sei, eine Wahrheit, die ausschließlich durch die reine Vernunft (ohne Überprüfung in der Erfahrung) feststellbar ist. Kants Ansicht zufolge sind wir Menschen nicht fähig, eine *nicht*kausale Welt zu erfahren oder über eine solche nachzudenken, denn unsere Urteile sind an die Gesetze der Natur und damit an das Prinzip der Kausalität gebunden. Anders formuliert: Unsere Urteile

haben bestimmte Formen und die Ordnung oder das Prinzip der Kausalität sind diesen Formen implizit. Wenn wir keine Urteile dieser Form treffen könnten, wären wir keine rational handelnden Wesen. Kausalität ist also keine Illusion, sie ist vielmehr Erkenntnis. Aus heutiger Sicht können wir Kants Aussage etwa folgendermaßen paraphrasieren: Kausalität ist eine der Erkenntnis eingeborene Ordnung, die wir auf die Welt anwenden, wenn wir bestimmte Arten von Ereignissen interpretieren. Sie ist ein natürlicher Teil unserer kognitiven Architektur.

Was verursacht was? – Wie Experten diese Frage beantworten

Diese Ideen der Philosophen des 18. Jahrhunderts sind bis heute in der psychologischen Literatur über kausale Kognition äußerst lebendig. Eine sehr einflussreiche neuzeitliche Betrachtung der kausalen Kognition wurde von Patricia Cheng (1997) und Laura Novick (Cheng & Novick 1992) vorgelegt. Cheng sagt: „Kausale Beziehungen sind weder unmittelbar beobachtbar noch ableitbar […] der logisch denkende Beobachter glaubt, dass es Dinge in der Welt wie Ursachen gibt, welche die Kraft haben, eine Wirkung zu erzeugen, und dass es Ursachen gibt, welche die Kraft haben, eine Wirkung zu verhindern, und dass eben nur solche Dinge das Auftreten einer Wirkung beeinflussen" (Cheng 1997). Aus psychologischer Sicht ist die Kausalität der Ereignisse also nicht beobachtbar, sondern wird von uns Menschen erschlossen. Aber für jedes beliebige Ereig-

nis gibt es eine Vielzahl möglicher Erklärungen (Ursachen/Gründe). Wenn Ihr Auto nicht anspringt, könnte es sein, dass die Batterie leer ist, dass der Tank leer ist oder dergleichen mehr. Doch wie ermitteln wir die wahren Ursachen eines Ereignisses?

Laut Cheng und Novick gehen wir nach dem sogenannten Covariationsprinzip vor, wenn mehrere Erklärungen (Ursachen) für eine Wirkung infrage kommen. Das heißt, wir suchen nach Faktoren, die als Ursache für ein Ereignis relevant sein können (Covariationsinformationen) und nehmen eine Bewertung derselben vor, um unter möglichen Ursachen auszuwählen und zu einer Erklärung zu gelangen. Dabei machen wir diejenige Ursache für ein Ereignis verantwortlich, die mit diesem Ereignis am stärksten *covariiert* oder die höchste statistische *Kontingenz* aufweist. Als formalisiertes Modell sieht dies folgendermaßen aus:

$$\Delta P = p(e \mid c) - p(e \mid -c)$$

e steht für Wirkung (Effekt)
c steht für eine (beliebige) Ursache
p(e | c) steht für die Wahrscheinlichkeit von e, vorausgesetzt c liegt vor
p(e | –c) steht für die Wahrscheinlichkeit von e, vorausgesetzt c liegt nicht vor

Verglichen wird dabei jeweils die Wahrscheinlichkeit des Effekts bei Vorliegen der Ursache – P(Effekt|Ursache) = P(e|c) – mit der Wahrscheinlichkeit des Effekts ohne Vorliegen der Ursache – P(Effekt|keine Ursache) = P(e|–c). Die

Differenz der beiden Wahrscheinlichkeiten wird dabei auch als *Kontingenz* oder ΔP bezeichnet. Ursache und Wirkung müssen also, wie Hume postulierte, zusammen auftreten. Zurück zu unserem „Auto"-Beispiel: Nehmen wir an, Sie haben gerade einen Gebrauchtwagen gekauft. Doch der hat, wie Sie im Nachhinein feststellen müssen, eine technische Macke, nämlich die, dass sich die Batterie entlädt, sobald das Radio zu lange läuft. Und so passiert es Ihnen des Öfteren, dass der Wagen nicht mehr anspringt, weil Sie während der letzten Fahrt zu lange Radio gehört haben und die Batterie nun leer ist. Entnervt stellen Sie fest: „Scheint, als würde die Batterie schlappmachen, jedes Mal, wenn ich Radio höre!" In einem solchen Fall wäre ΔP sehr hoch, denn die Wahrscheinlichkeit, dass die beiden Ereignisse – Radio hören und leere Batterie – zusammen auftreten, ist größer als die, kein Radio zu hören und trotzdem eine leere Batterie zu haben.

Nach Cheng und Novick können Ursachen also aufgrund der probabilistischen Zusammenhänge erkannt werden: Wenn ΔP erkennbar positiv ist (und in seinem Wert über einem bestimmten Kriterium liegt), dann wird c als eine *generative* (erzeugende/fördernde) Ursache erkannt. Wenn ΔP jedoch negativ ist, ist c eine *inhibitorische* (verhindernde) Ursache. Und wenn ΔP positiv ist, aber unter dem bestimmten Kriterium liegt, wird c als nichtkausal bewertet. Cheng (1997) führte ihre Analyse noch einen Schritt weiter und entwickelte eine Formel, die kausale Ereignisse und Koinzidenzen (das Zusammentreffen von Ereignissen) noch besser unterscheidet:

$$p_c = \Delta P / [1 - p(e \mid \sim c)]$$

Dieses Modell vergleicht die kausale Stärke einer behaupteten Ursache mit der Wahrscheinlichkeit des Effekts, der unter Vorliegen anderer Ursachen auftritt. Fügen wir nun ein paar Zahlen ein, um zu sehen, wie es funktioniert. Angenommen, Sie haben 20 Fahrten gemacht. Dabei haben Sie auf 10 Fahrten Radio gehört und auf 10 nicht. Acht von den 10 Fahrten, die Sie Radio gehört haben, war die Batterie danach leer. Die Wahrscheinlichkeit, dass die Batterie unter diesen gegebenen Umständen leerläuft, stellt sich demnach so dar: 8/10=0,8. In 80 % dieser Fälle hat die Batterie also ihren Geist aufgegeben. Wie sieht es aber mit den anderen 10 Fahrten aus, die Sie das Auto benutzt haben, ohne Radio zu hören? Nehmen wir einmal an, die Batterie war danach in 2 von 10 Fällen leer. Die Wahrscheinlichkeit für diesen Fall stellt sich so dar: 2/10=0,2 (oder 20 %). Und nun können wir ΔP ganz einfach berechnen: 0,8 – 0,2=0,6. Vorausgesetzt ΔP reicht von –1 bis 1, wobei 0 kein Kausalzusammenhang bedeutet, deutet ein Wert von 0,6 sehr wahrscheinlich darauf, dass ein generativer kausaler Zusammenhang besteht (Radio laufen lassen bewirkt, dass die Batterie leerläuft), ist aber noch kein besonders stichhaltiger Beweis. Betrachten wir nun die Fälle, in denen Sie sich mit einer leeren Batterie herumplagen mussten, obwohl Sie das Radio gar nicht an hatten. Dies bedeutet, es gibt alternative Ursachen (wie ein Fehler in der Batterie selbst). Der Nenner in der obigen Gleichung benennt die Wahrscheinlichkeit, mit der der Effekt (leere Batterie) eher durch das Radio als durch eine alternative Ursache verursacht wurde. Der Ausdruck p(e|~c) stellt die Wahrschein-

lichkeit dar, dass die Wirkung ohne Vorliegen der Ursache auftritt – das heißt, eine leere Batterie, auch ohne dass das Radio lief. Genau das hatten wir zuvor berechnet: 2 von 10 oder 0,2. Subtrahieren wir nun diesen Wert von 1, haben wir das Verhältnis der Häufigkeiten, mit der die Wirkung auftrat, ohne dass die Ursache vorlag. Das ist ebenfalls leicht zu rechnen: $1 - 0{,}2 = 0{,}8$. Wenn wir ΔP nun durch diesen Wert teilen, bekommen wir p_c. Und p_c sagt uns, wie viel Stärke der Kausalbeziehung zwischen Radio und leerer Batterie zugeschrieben werden kann im Vergleich zu anderen möglichen Ursachen. Dieser Wert stellt sich für unser Beispiel wie folgt dar: $0{,}6/0{,}8 = 0{,}75$. Unter der Voraussetzung, dass p_c von 0 bis 1 reichen kann, ist dies ein stärkerer Beweis dafür, dass das Radio der Übeltäter ist. Das Radio als Ursache für die leere Batterie erhielt erfahrungsbasiert einen moderaten Wahrscheinlichkeitswert (0,6), erwies sich aber im Vergleich mit alternativen Ursachen als der wahrscheinlichste Kandidat (0,75) für einen kausalen Einfluss.[1]

Einen Haken hat die Sache allerdings: Das Modell von Cheng und Novick sagt uns zwar, wie wir unter potenziellen Erklärungen (Ursachen) für ein Ereignis zu einem Schluss gelangen, nämlich indem wir sogenannte Covariationsinformationen verwenden (das heißt, indem wir nach Faktoren suchen, die als Ursache für ein Ereignis relevant

[1] Der Mathematiker Judea Pearl (2000) hat gezeigt, dass die von Patricia Cheng vorgestellte Kausaltheorie auch kontrafaktisch interpretiert werden kann (d. h. unter der Wahrscheinlichkeit, dass, vorausgesetzt c und e hätten nicht stattgefunden, e wahr ist, wenn c wahr ist. Die kontrafaktische Implikation besteht darin, dass Chengs Modell unter Verwendung von Strukturmodellen berechenbar ist. Griffiths und Tenenbaum (2005) zeigten des Weiteren, dass Chengs Modell als eine verrauschte, nichtinklusive Oder-Funktion (nichtausschließende Disjunktion) interpretiert werden kann, um Wahrscheinlichkeiten zu berechnen.

sein können). Aber das ist, wie sich gleich herausstellt, nur die halbe Geschichte.

Das Problem lässt sich denkbar einfach veranschaulichen: Wir bleiben bei unseren obigen Zahlen, stellen uns nun aber vor, dass wir bei 8 von 10 Fahrten, nach denen die Batterie leer war, nicht Radio gehört, sondern ein belegtes Brötchen als Pausensnack zur Arbeit mitgenommen haben. Unsere Werte für ΔP und p_c sind genau gleich. Doch es fällt Ihnen ungleich schwerer, sich vorzustellen, dass ein belegtes Brötchen zu einer leeren Batterie führen könnte. Doch das, so schiebt Cheng nach, liege daran, dass ein belegtes Brötchen in Ihrem Spektrum der potenziellen kausalen Topkandidaten gar nicht vorkomme. Drängt sich bloß die Frage auf, warum es nicht vorkommt.

Kausaltheoretiker nehmen diesen Einwand sehr ernst. Sie heben vor allem darauf ab, dass Menschen kausale Urteile treffen (oder Schlussfolgerungen ziehen), indem sie nach Informationen über mögliche generative kausale Mechanismen suchen. Ahn et al. (1995) drücken dies folgendermaßen aus: „[…] ein Mechanismus ist eine Komponente eines Ereignisses, von dem man glaubt, es wohne ihm eine kausale Kraft oder kausale Notwendigkeit inne […] liegen zwei kausal zusammenhängende Ereignisse zugrunde, gibt es ein System von verbundenen Teilen, die wirken oder aufeinander einwirken, um ein Ereignis hervorzubringen oder zu erzwingen." Um dies zu demonstrieren, legten sie Probanden Beschreibungen kausaler Ereignisse vor wie etwa: „Kim hatte gestern Abend einen Autounfall", gefolgt von Aussagen, die eine Mechanismusinformation gaben (z. B. „Kim ist kurzsichtig und hat beim Autofahren meist ihre Brille

nicht auf"), oder eine Covariationsinformation (z. B. „Gestern Abend war mit mehr Verkehrsunfällen zu rechnen"). Die Probanden sollten nun abschätzen, in welchem Grad die genannte Information (Ursache) verantwortlich war für das aufgetretene Ereignis. Wie die abgegebenen Einschätzungen zeigten, wurde die Mechanismusinformation ungefähr doppelt so stark bewertet wie die Covariationsinformation. Die Ergebnisse waren erstaunlich: Die Probanden zogen nur selten die Covariationsinformation heran, wenn sie gebeten wurden, die Ursache eines Ereignisses zu bestimmen, auch wenn diese Informationen ohne Weiteres zugänglich waren. Stattdessen waren ihre Überlegungen theoriebasiert und zielten darauf ab, einen Mechanismus aufzudecken, der das Ereignis hervorbringen könnte. Selbst wenn Covariationsinformationen verlangt und vorgegeben waren, begründeten die Probanden ihre Attribution nicht mit Bezug auf die Covariationsinformation. Stattdessen erklärten sie die Kausalzusammenhänge mehrheitlich mechanismus- und nicht covariationsbasiert. White fand außerdem heraus (1995, 2000), dass selbst wenn ein „kausaler Kandidat" perfekt mit einer Wirkung korrelierte, er nicht als eine Ursache identifiziert wurde, wenn ihm nicht auch die Kraft zugeschrieben wurde, diese Wirkung zu erzeugen. Es scheint also tatsächlich so, als beharrten wir auf jene „notwendige Verknüpfung", die bereits Hume so problematisch fand.

Um dies noch deutlicher zu machen, betrachten wir ein einfaches Beispiel: Angenommen die Häufigkeit des Auftretens von Autounfällen korreliert vollkommen mit der Häufigkeit von a) Tätowierungen der Autofahrer und b)

einem neuen Bremssystem des Autos. Die meisten werden sich wohl für b) als wahre Ursache entscheiden und a) völlig verwerfen – es sei denn, sie sind der Meinung, die Tätowierungen covariierten mit einem weiteren Faktor, der als generative Ursache für dieses Ereignis relevant sein könnte (wie etwa dem Klischee, dass Menschen mit einer Tätowierung sich eher unter dem Einfluss von Drogen oder Alkohol hinters Steuer setzen – oder dass die Tätowierungen toxische Substanzen in den Körper abgeben, die das Urteilsvermögen beeinträchtigen).

Dabei ist es gar nicht mal so verkehrt, eine gesunde Skepsis walten zu lassen und unser Urteil nicht allzu sehr auf Covariationen zu stützen, wenn wir kausale Rückschlüsse ziehen. Judea Pearl (2009), Informatiker und Philosoph, der zu den führenden Experten auf dem Gebiet der Kausalmodellierungen gehört, behauptet, dass es sinn- und zwecklos sei, Kausalität allein aufgrund der probabilistischen Relationen zwischen Ereignissen zu definieren. Wo dies geschieht, so Pearl, stelle sich bei näherer Betrachtung heraus, dass unsere Urteile meist auf versteckten kausalen Annahmen beruhen, die sich sehr anschaulich in kontrafaktischen oder mechanistischen Annahmen äußern („wenn dies oder jenes nicht passiert wäre, wäre auch das oder das nicht passiert"). Und dies wiederum führe bisweilen zu Verzerrungen im logisch kausalen Denken, da wir objektive Daten, die eine kausale Theorie stützen, ausblenden, weil wir die Theorie nicht verstehen oder sie als störend empfinden. Ein aktuelles Beispiel hierfür ist die Weigerung einer großen Mehrheit der US-Amerikaner, die Evolutionstheorie anzuerkennen, trotz umfassender Beweise, die ihre Gültigkeit stützen; laut einer

Umfrage des US-amerikanischen Meinungsforschungsinstituts Gallup Organization von 2009 erkennen nur vier von zehn US-Bürgern die Evolutionstheorie an.

Was verursacht was? – Wie unser Gehirn diese Frage beantwortet

Theoriebasiertes kausales Denken ist eine Art annahmebasiertes Denken. Wir sind der vorgefassten Meinung, dass kausalen Beziehungen Mechanismen zugrunde liegen, die bewirken können, dass ein Ereignis ein anderes hervorbringt. Der Einfluss vorgefasster Meinungen auf die kausale Kognition wurde in einer neurowissenschaftlichen Studie, durchgeführt von Fugelsang und Dunbar (2005), in drastischer Weise deutlich. Im Rahmen dieser Studie waren die Probanden aufgefordert, kausale Szenarien zu bewerten, während eine funktionelle Magnetresonanztomografie durchgeführt wurde. Die Probanden bekamen jeweils vier Szenarien in Folge gezeigt, zwei, die plausible kausale Szenarien beschrieben, und zwei, die nichtplausible kausale Szenarien beschrieben. Während der „plausiblen Szenarien" sagte man ihnen, dass ein höherer Serotoninspiegel die Stimmung aufhelle (plausibel), und eine rote Pille wurde als Arzneimittel bezeichnet, das den Serotoninspiegel erhöht. Während der „nichtplausiblen Szenarien" wurde ihnen mitgeteilt, Antibiotika hätten keinerlei Auswirkungen auf die Stimmung, und eine rote Pille wurde als Antibiotikum bezeichnet. Anschließend wurde den Probanden eine zufällige Reihe von Zeichnungen präsentiert, die immer die rote

Pille mit einem glücklichen oder einem traurigen Gesicht sowie eine blaue Pille (Placebo) mit einem glücklichen oder einem traurigen Gesicht verbanden. Das Verhältnis der Covariationen lag entweder bei 18 von 22 ($\Delta P = 0{,}74$) und damit bei einem mäßigen bis starken kausalen Zusammenhang, oder bei 10 von 22 ($\Delta P = 0{,}30$) und damit bei einem vergleichsweisen schwachen kausalen Zusammenhang. Abschließend wurden die Probanden gebeten, anhand einer Drei-Punkte-Skala (niedrig, mittel, hoch) zu bewerten, wie stark sich die rote Pille auf eine Erhöhung der Glücksgefühle ausgewirkt hatte.

Die Ergebnisse überraschten: Die Hirnareale, die mit Lernen assoziiert sind, waren am aktivsten, wenn Theorie und (probabilistische) Informationen konsistent waren (plausible und starke Covariation, oder nichtplausible, schwache Covariation), insbesondere, wenn die Probanden (probabilistische) Informationen bewerteten, die mit einer plausiblen Theorie konsistent waren. Die Hirnareale, die mit Denken und Aufmerksamkeit assoziiert sind, waren am aktivsten, wenn probabilistische Informationen (Daten) und Theorie inkonsistent waren (nichtplausible Theorie mit starker Covariation oder plausible Theorie mit schwacher Covariation), insbesondere wenn plausible Theorien schwach unterstützt waren. Was bedeutet dies? Wir befassen uns intensiver mit der Verarbeitung von probabilistischen Informationen, die mit unseren Annahmen aus diesen Informationen nicht konsistent sind, lernen aber nicht zwangsläufig daraus (insofern, als dass wir unsere Annahme revidieren würden). Die Autoren folgerten, dass es dadurch eine starke Verzerrung im kausalen Denken gibt, die sich

darin zeigt, dass wir uns erstens auf Theorien fokussieren, die mit unseren Annahmen konsistent sind, zweitens mit Informationen aufhalten, die mit unseren Meinungen nicht konsistent sind, und drittens wir unsere Annahmen nicht zwangsläufig revidieren. Das Denken in kausalen Zusammenhängen ist also nicht gerade unsere Stärke. Wir weigern uns entschieden (im wortwörtlichen Sinne!), unsere Annahme zu ändern, selbst angesichts zwingender Beweise.

Was sind notwendige, was hinreichende Bedingungen?

Diese Ergebnisse zeigen, dass kausales Denken typischerweise nicht vonstattengeht, ohne dass ein bestimmtes Vorwissen oder vorgefasste Annahmen abgefragt und herangezogen würden. Wir können unsere Frage also leicht abwandeln: Inwiefern beeinflussen bereits vorhandene Kenntnisse und Annahmen kausale Folgerungen? Betrachten wir als Beispiel folgende Aussage: „In einen Swimmingpool voll Wasser zu springen bewirkt, dass die Kleider nass werden" – eine vernünftige Annahme, gefasst aufgrund jeder Menge beobachteter Covariationen dieses Ereignisses: Wir wissen, dass der Sprung in einen Swimmingpool voll Wasser in der Tat die kausale Kraft hat, Kleider nass zu machen. Nehmen wir nun an, Sie sind auf eine Poolparty eingeladen, machen sich auf den Weg dorthin und kurz bevor Sie an der Haustür klingeln, geht diese auf und eine Person mit nassen Kleidern am Leib kommt heraus. Sie waren schon auf vielen Poolpartys, daher wissen Sie, dass eine solche Par-

ty eine recht nasse Angelegenheit ist, es ordentlich platscht und spritzt, wenn die Gäste ins Becken springen und ausgelassen im Wasser tollen, sodass man auch am Beckenrand neben dem Pool oft ziemlich nass wird. Aber folgern wir deshalb, dass die Person, die wir eben aus der Tür kommen sahen, im Pool, also im Wasser war?

Natürlich nicht. Unter diesen Umständen (*Bedingungen*) würden wir diesen Schluss wahrscheinlich nicht ziehen, weil wir unser Erfahrungswissen abrufen und annehmen, dass es neben einem Sprung ins feuchte Nass auch andere Gründe geben könnte, die erklären, warum die Kleider dieser Person gerade nass sind. Und weil es viele alternative Ursachen geben kann, denken wir, dass „in den Pool springen" kein *notwendiger Grund* für die nassen Kleider ist. Wenn wir jedoch mit eigenen Augen gesehen hätten, wie diese Person in voller Montur ins Wasser gesprungen ist, hätten wir uns wohl sehr gewundert, denn zwischen dem Ursachenfaktor („ins Wasser springen") und dem Effekt („nasse Kleider") scheint *hinreichender Zusammenhang* zu bestehen.

Vergleichen wir nun diese Situation mit der folgenden: Sie gehen mit einer Freundin zum Kanufahren und Camping. Am Abend sehen Sie ihr dabei zu, wie sie versucht, ein Lagerfeuer zu machen. Sie häuft Blätter, Zweige und trockenes Gras auf und zündet dann ein Streichholz, um das Feuer zu entfachen. Würde das Streichholz nicht direkt zünden, würden Sie sich wohl nicht gleich wundern, denn normalerweise, und das wissen Sie, funktioniert es. Sie ziehen den kausalen Zusammenhang zwischen „Streichholz anzünden" und „Feuer entfachen" nicht in Zweifel. Es gibt allerdings Faktoren, die unter den gegebenen Bedingungen

hinzukommen und eben diesen Effekt verhindern können. Das Streichholz könnte bei der Kanufahrt nass geworden sein. Doch das bedeutet nicht, dass Sie den kausalen Zusammenhang zwischen Streichholz und Feuer widerrufen. Es bedeutet vielmehr, dass die Ursache zwar vorliegt („Streichholz entzünden"), aber nicht *hinreichend* ist für den Effekt („Feuer entfachen"). Der Ursache-Wirkungs-Zusammenhang wurde durch das Wasser auf dem Streichholz verhindert.

Damit haben wir zwei neue, wichtige Begriff eingeführt: *kausal notwendige Bedingung* und *kausal hinreichende Bedingung*. Diese beiden Begriffe haben in der Philosophie eine lange Geschichte. Schon 1748 definiert der schottische Empirist David Hume das Wesen der Ursache und Kausalität als „einen Gegenstand, dem ein anderer folgt, wobei allen Gegenständen, die dem ersten gleichartig sind, Gegenstände folgen, die dem zweiten gleichartig sind. Oder mit anderen Worten: wobei, wenn der erste Gegenstand nicht bestanden hätte, der zweite nie ins Dasein getreten wäre".[2]

Nach dieser Definition ist eine Ursache eine notwendige Bedingung für das Auftreten eines bestimmten Ereignisses. Man beachte, dass Hume die kausale Notwendigkeit als kontrafaktisch bekundet – wenn die Ursache nicht aufgetreten wäre, wäre auch der Effekt nicht eingetreten. Auf diesen kontrafaktischen Konditionalen baut auch die Theorie des zeitgenössischen Philosophen David Lewis auf (1973, 1979), in der Hume einen Nachhall findet. Bei Lewis heißt

[2] Kulenkampff J (1993) Eine Untersuchung über den menschlichen Verstand. Übersetzt von Raoul Richter. 12. Aufl. Meiner, Hamburg, S. 92f

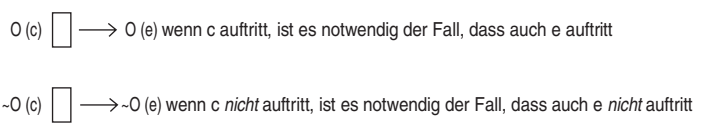


Abb. 6.1. Die kontrafaktischen Konditionalen der Kausalität nach dem Philosophen David Lewis

es: „Wo C und E als aktuale Ereignisse vorliegen, heißt ein Einzelereignis E von einem Einzelereignis C kausal abhängig, wenn E nicht eingetreten wäre, wäre auch C nicht eingetreten". Diese Beziehungen lassen sich in der Modallogik Modallogik auch formalisiert darstellen, wobei das Kastensymbol die *notwendige Ursache* darstellen soll. O(e) soll alle und nur jene möglichen Welten darstellen, in denen e (der Effekt) eintritt; O(c) soll alle und nur jene möglichen Welten darstellen, in denen c (die Ursache) vorliegt. Danach ist ein Einzelereignis E von einem Einzelereignis C kausal abhängig, wenn die in Abb. 6.1 dargestellten kontrafaktischen Bedingungen gelten.

Warum stellt Lewis dieser Analyse den Satz „Wo C und E als aktuale Ereignisse vorliegen" voran? Die Antwort darauf lautet: Weil er argumentiert, dass Nicht-Ereignisse (das Nicht-Eintreten von Ereignissen) nicht als Ursache (für andere Ereignisse) dienen können. Ein Beispiel: Mal angenommen, Sie beobachten jemanden dabei, nennen wir ihn Joe, der über eine Leiter auf das Dach eines Hauses steigt. Der Satz „Dass Joe die Leiter besteigt, bewirkt, dass er aufs Dach gelangt" beschreibt zwei aktuale Ereignisse: Das Auftreten des einen (Joe steigt die Leiter hinauf) ist abhängig

vom Auftreten des anderen (Joe gelangt aufs Dach). Der Satz „Die Leiter, die nicht gebrochen ist, bewirkt, dass Joe aufs Dach gelangt" beinhaltet keine kausale Behauptung (trotz der kontrafaktischen Abhängigkeit, die er impliziert), denn die Leiter, die nicht bricht, ist ein Nicht-Ereignis. Die Tatsache, dass wir Nicht-Ereignisse oft einen kausalen Einfluss zuschreiben, führte Lewis zu der viel extremeren Behauptung, wonach die Ursache-Wirkungs-Relation nicht als grundlegend für das Phänomen der Kausalität anzusehen sei.

Der britische Empirist John Stuart Mill (1843) brachte eine explizite Analyse der beiden Begriffe *kausal notwendig* und *kausal hinreichend* vor. Nach Mill ist eine *kausal notwendige* Ursache für das Auftreten von e ein Faktor, der in allen (möglichen) Fällen eines Effekts e vorliegt; eine *kausal hinreichende* Ursache für das Auftreten von e ist ein Faktor, der das Auftreten eines Effekts e garantiert. Ein „Faktor" kann ein Einzelereignis sein (Batterie läuft leer), eine Eigenschaft (Batterie ist vollständig aufgeladen) oder eine Variable (wie die Temperatur). Der Faktor, der den größten Unterschied zwischen dem kausalen Einfluss und dem Auftreten von e ausmacht, ist der zentrale Faktor (für unser Beispiel, die Batterie läuft leer). Mill schlug das *deterministische Prinzip* vor, das besagt, dass auf exakt gleiche Ursachen immer exakt gleiche Wirkungen folgen (eine leere Batterie führt immer dazu, dass das Auto nicht anspringt). Mills Grundgedanken hierzu werden zusammengenommen als *Mills Kanones* bezeichnet.

Der australische Philosoph John Mackie führte diese Gedanken weiter, indem er folgende Analyse vorstellt:

- *hinreichend* bedeutet, dass eine Ursache durch sich selbst eine Wirkung erzeugen kann (wenn c, dann e)
- *notwendig* bedeutet, dass eine bestimmte Ursache vorliegen muss, damit eine (bestimmte) Wirkung eintreten kann (wenn nicht c, dann nicht e)
- ein *notwendiger und hinreichender* kausaler Zusammenhang besteht, wenn nur eine einzige Ursache für eine Wirkung vorliegt (wenn und nur wenn c, dann e)

Da ein Ereignis meist nicht durch ein einziges Ereignis als hinreichende Bedingung verursacht wird, sondern oft mehrere Ursachen hat, führte Mackie 1974 die sogenannte INUS-Bedingung ein (*insufficient, but necessary part of an unnecessary but sufficient condition*). Danach ist eine Ursache ein *hinreichender*, aber nicht *notwendiger* Teil einer Bedingung, die wiederum nicht *notwendig*, aber *hinreichend* ist für das Auftreten des Ereignisses (Wirkung).

Wie wichtig es ist, zwischen *kausal notwendigen* und *kausal hinreichenden* Bedingungen zu unterscheiden, führen wissenschaftliche Erklärungen vielleicht am deutlichsten vor Augen. Die Philosophen Nancy Cartwright (1980) und David Armstrong (1983) argumentieren, naturgesetzliche Zusammenhänge in der wirklichen Welt seien eher „eichern" denn „eisern", „eicherne" Gesetze nämlich ließen Ausnahmen zu: Gesetzmäßige Zusammenhänge zwischen verschiedenen Ereignissen hängen von impliziten *ceteris paribus*-Bedingungen ab (das heißt „unter sonst gleichen Bedingungen") oder von *ceteris absentibus*-Bedingungen (das heißt „unter sonst fehlenden Bedingungen). Aus diesem Grund ist eine Folgerung, die auf einem Naturgesetz basiert, immer anfechtbar, da sich jederzeit herausstellen

kann, dass wir dem infrage stehenden Gesetz im besonderen Fall zusätzliche Faktoren hinzufügen müssen.

Wie unsere Überzeugungen unser Entscheidungsverhalten beeinflussen

Im vorangegangenen Kapitel haben wir davon gesprochen, dass unser alltägliches Denken sehr stark geprägt ist von spezifischen Informationen oder Ausnahmeninformationen. Wir nutzen diese Art der Information, um zwischen *kausal notwendigen* und *kausal hinreichenden* Bedingungen zu unterscheiden und dann ein Kausalurteil abzuleiten.

Insbesondere Cummins et al. haben gezeigt, dass wir bei der Bewertung kausaler Zusammenhänge unseren Gedächtnisspeicher nach vielen alternativen Ursachen (*disablers*) durchforsten, die Informationen auslesen und abrufen (Cummins 1995, 1997; Cummins et al. 1991). Im Rahmen zahlreicher Studien sollten die Probanden entscheiden, ob ihrer Meinung nach ein bestimmtes Ereignis auftreten wird oder einem bestimmten Ereignis eine kausale Rolle zuzuschreiben ist. Wie sich zeigte, suchten die Probanden in ihrer Urteilsfindung nach alternativen Ursachen ebenso wie nach möglichen *disablers*, die verhindern können, dass ein bestimmtes Ereignis auftritt, auch wenn ein plausibler Grund für das Auftreten des Ereignisses vorliegt. Alternative Ursachen führen dazu, dass wir *kausal notwendige* Verknüpfungen bezweifeln, und *disablers* führen dazu, dass wir *kausal hinreichende* Verknüpfungen bezweifeln.

Disablers sind von großer Bedeutung für unser kausales Denken, denn sie provozieren einschreitende Handlungen (oder legen diese nahe), um unerwünschte Ereignisse abzuwenden. Zum Beispiel ist weithin bekannt, dass misshandelte Kinder ein höheres Risiko tragen, später selbst zu misshandelnden Eltern zu werden. Doch dieser Zusammenhang ist alles andere als gewiss. Vielmehr deuten Forschungen darauf, dass mehrere Faktoren zu möglichen Missbrauchspraktiken in der Kindererziehung führen können, aber das Vorliegen eines nichtmisshandelnden elterlichen Rollenvorbilds in der Kindheit mindert die Wahrscheinlichkeit, dass ein missbrauchtes Kind später selbst zu einem misshandelnden Elternteil wird (Kaufman & Zigler 1988; Martin & Elmer 1992). Nichtmisshandelnde Rollenvorbilder sind deshalb *disablers*, die das Auftreten eines Effekts verhindern können (künftige misshandelnde Eltern), und zwar trotz Vorliegen einer tatsächlichen Ursache (misshandelt zu werden).

Disablers heben einen kausalen Zusammenhang nicht auf. Kindesmisshandlung selbst erfahren zu haben, bleibt eine tatsächliche Ursache für selbsttätige Misshandlungen in der späteren Elternrolle, aber die Ursache-Wirkungs-Relation besteht nicht zwangsläufig. Sie kann unterbrochen oder verhindert werden, indem ursächliche Faktoren entschärft werden, zum Beispiel dadurch, dass ein nichtmisshandelndes Rollenvorbild vorhanden ist. Dies ist die Umkehrung *günstiger Bedingungen* – Hintergrundfaktoren, die vorhanden sein müssen, damit eine tatsächliche Ursache auch tatsächlich einen bestimmten Effekt erzeugt, wenngleich sie selbst nicht die Ursachen sind. Beispiel: Das Vorhandensein von Sauerstoff ermöglicht Verbrennungsre-

aktionen bei gleichzeitigem Auftreten einer tatsächlichen Ursache, wie etwa dem Entzünden eines Streichholzes. *Disablers* sind Hintergrundfaktoren, die nicht vorhanden sein müssen, damit eine tatsächlich vorhandene Ursache ein bestimmtes Ereignis hervorbringt. Unter Verwendung von Chengs Terminologie können wir sie als *präventive* Ursachen betrachten (obgleich ich diesen Ausdruck etwas sperrig finde.)

Die Aktivierung von Erfahrungswissen auf der Suche nach alternativen Ursachen und *disablers* ist ein wichtiger Teil kausaler Denkprozesse. Diese Faktoren nämlich beeinflussen das logische Folgern und Urteilen, auch wenn sie darin nicht explizit auszumachen sind. Cummins et al. fanden heraus, dass die Probanden nur zögerlich von einer bestimmten Ursache auf eine Wirkung schlossen, wenn weitere alternative Ursachen möglich waren. Und sie taten sich offenbar schwer, eine Ursache als eine hinreichende Bedingung für einen bestimmten Effekt anzusehen, wenn weitere *disablers* möglich waren. Damit scheint sich zu bestätigen, dass wir *kausal notwendige* Verknüpfungen bezweifeln, wenn alternative Ursachen vorliegen, während wir *kausal hinreichende* Verknüpfungen bezweifeln, wenn *disablers* vorliegen. Dieser Effekt zeigte sich in mehrfacher Wiederholung bei Erwachsenen (z. B. de Neys et al. 2002, 2003) und Kindern (z. B. Janveau-Brennan & Markovits 1999).

Im Rahmen einer weiteren Studie, wiesen Verscheuren et al. (2003) ihre Probanden an, während ihrer kausalen Urteilsprozesse laut zu denken. Wie sich herausstellte, zogen die Probanden spontan alternative Ursachen heran, wenn es zu entscheiden galt, ob eine bestimmte Ursache *notwendig* sei, um einen bestimmten Effekt zu erzeugen.

Und sie zogen spontan *disablers* heran, wenn es zu entscheiden galt, ob eine bestimmte Ursache *hinreichend* sei, um einen bestimmten Effekt zu erzeugen. Je größer die Zahl der herangezogenen alternativen Faktoren, desto langsamer fanden sie zu einer Entscheidung (de Neys et al. 2003).

Das Paradox der Kausalität – Wie sich Kinder die Welt erschließen

Ich möchte dieses Kapitel gerne beschließen, indem ich noch einmal zurückkomme auf das eingangs angesprochene Paradox der Kausalität. Wie können wir über Kausalität nachdenken, wenn sie sich unserer unmittelbaren Wahrnehmung entzieht? Eine Antwort auf diese Frage kommt aus der Forschung über frühkindliche kognitive Prozesse zum Wissenserwerb. Die wohl wichtigste Erkenntnis aus den Forschungsergebnissen der vergangenen Jahrzehnte ist die, dass ein großer Teil des Grundwissens, das wir brauchen, um uns die Welt sinnhaft zu erschließen, entweder von Geburt an vorhanden ist oder in früher Kindheit sehr schnell entsteht – zu schnell, um es durch Lernerfahrungen wie Versuch-und-Irrtum erst mühsam zu erwerben. Ein Aspekt aber ist für die kognitive Strukturbildung besonders wichtig. Es ist das Verständnis von Kausalität, das sich innerhalb der ersten sechs bis acht Lebensmonate entwickelt.

Erste Experimente, die die Entwicklung des logisch-kausalen Denkens bei Kindern untersuchten, nutzten eine Methodik, die Albert Michotte , einem belgischen Experimentalpsychologen, zu verdanken ist. Nach Michotte (1963)

ist die wichtigste Voraussetzung für die Wahrnehmung von kausalen Zusammenhängen die Übertragung von Bewegung, Energie, Schwungkraft oder Bewegungskraft. Der einfachste zu beobachtende Kausalzusammenhang besteht in kausalen Bewegungsereignissen, in der sichtbaren Übertragung von Energie von einem Objekt auf ein anderes: Ein Objekt wird angestoßen, kollidiert mit einem zweiten Objekt und setzt dieses in Bewegung. Michotte ging davon aus, dass uns die Fähigkeit, solche Ereignisse als kausale Ereignisse auszumachen und zu begreifen, angeboren sei und dass sie die Grundlage bildet für alle kausalen Erkenntnisse, die wir später hinzugewinnen.

Michotte überprüfte seine Theorien nicht an Kindern, aber er konzipierte etliche Versuchsreihen, die er an Erwachsenen testete. Seine systematischen Versuche sind simpel, aber sehr anschaulich und überzeugend: Die Teilnehmer beobachten auf einer Leinwand die Projektion zweier einfacher Objekte (zum Beispiel eines roten und eines blauen Balls), die auf unterschiedliche Arten miteinander interagierten. In der Versuchsanordnung „direkter Bewegungszusammenhang" beispielsweise bewegt sich der blaue Ball in direkter Linie über die Leinwand auf einen unbewegten roten Ball zu, stößt ihn an und setzt ihn in Bewegung, sodass er in die gleiche Richtung davon rollt. Die meisten Erwachsenen nehmen dieses „Bewegungsereignis" als ein kausales Ereignis wahr – das heißt, der blaue Ball stößt den roten Ball an, wodurch der rote Ball in Bewegung gesetzt wird. In einer zweiten Versuchsanordnung, in einem „zeitlich verzögerten Zusammenhang", wandert der blaue Ball über die Leinwand, berührt den roten Ball, aber nun bewegt sich der rote Ball zeitlich leicht verzögert. Die meisten

Erwachsenen nehmen dieses Bewegungsereignis nicht als ein kausales Ereignis wahr. Und in einer dritten Versuchsanordnung, in einem „räumlich auseinanderliegenden Zusammenhang", wandert der blaue Ball über die Leinwand, hält dann aber kurz bevor er den roten Ball berührt an. Dann setzt sich der rote Ball in Bewegung. Die meisten Erwachsenen nehmen auch dieses Ereignis nicht als ein kausales Ereignis wahr, was daran liegen mag, dass es keinen direkten Kontakt zwischen beiden Objekten gegeben hat. Stattdessen sieht es so aus, als würde sich der rote Ball von selbst in Bewegung setzen. Michotte war überzeugt, dass die Wahrnehmung von Bewegungsereignissen, die auf einer raumzeitlichen Koordination basierten, der einzige Bestimmungsgrund des kausal logischen Schließens sei.

Ungefähr zehn Jahre später begannen Entwicklungspsychologen Michottes theoretische Postulate auf den Prüfstand zu stellen, indem sie seine Versuche mit Kindern unterschiedlichen Alters wiederholten. Die Kinder bekamen ebenfalls zwei sich nacheinander bewegende, auf eine Leinwand projizierte Gegenstände zu sehen, während die Forscher sich der Methode des Habituationsparadigmas bedienten, um die kindliche Aufmerksamkeit zu prüfen (z. B. Ball 1973). Diese Methode ist recht einfach: Zunächst wurde den Kindern ein und dasselbe Ereignis wiederholt präsentiert, bis sie davon gelangweilt waren und den Blick abwendeten (die Aufmerksamkeitsreaktion auf ein Ereignis geht im Vergleich zur erstmaligen Präsentation des Ereignisses um bis zu 50 % zurück). Anschließend wurde der Versuchsaufbau systematisch relevant verändert und die Dauer der Zeit, die die Kinder das nun neue Bewegungsereignis verfolgten, wurde aufgezeichnet. Wenn sie

die Aufmerksamkeit erneut darauf lenkten (sprich, wenn die Betrachtungsdauer signifikant anstieg), konnten sie den Unterschied zwischen dem alten und dem neuen Bewegungsereignis benennen. Ebenso wie die Erwachsenen bekamen die Kinder einen blauen Ball zu sehen, der sich immer wieder über die Leinwand bewegte und mit einem roten Ball interagierte. Auch die Versuchsabfolge war dieselbe: direkter Bewegungszusammenhang, zeitlich verzögerter Zusammenhang, räumlich auseinanderliegender Zusammenhang. Da also auf ein Bewegungsereignis immer wieder ein neues folgte, müsste das aufmerksame Interesse der Kinder in jedem dieser Fälle eigentlich neu erwachen. Doch das tut es nicht. Vielmehr fanden sie die beiden letzteren Bewegungsereignisse weit spannender als das direkte Bewegungsereignis. Was bedeutet das? Nun, offenbar legt diese Beobachtung nahe, dass Kinder bereits über ein Grundwissen kausaler Zusammenhänge verfügen, weshalb sie ein Ereignis, das den Gesetzen der Kausalität gehorcht, weniger spannend finden als eines, das gegen diese Gesetze zu verstoßen scheint – wie zum Beispiel, wenn der blaue Ball mit dem roten Ball zusammenstößt, der rote Ball sich aber dennoch nicht bewegt; und so staunen die Kinder nicht schlecht, wenn die Ursache zwar vorliegt, der erwartet Effekt aber nicht eintritt (Kotovsky & Baillargeon 2000).

Einige Studien waren so konzipiert, dass das aktuale Bewegungsereignis versteckt (nicht sichtbar) ablief: Zum Beispiel war auf der Leinwand ein blauer Ball zu sehen, eine Trennwand und ein roter Ball, der halb versteckt hinter dieser Trennwand hervorlugte. Nun rollte der blaue Ball heran, rollte hinter die Trennwand und schon rollte der rote Ball hinter der Trennwand hervor. Die Kinder wurden also

daran gehindert, das kausale Ereignis mit eigenen Augen zu sehen, und mussten daher folgern, dass der blaue Ball den roten Ball berührt und angestoßen hat. Und das taten die Kinder auch. Hob man nun aber die Trennwand an und sie sahen, dass der blaue Ball den roten Ball gar nicht angestoßen hatte, dieser aber trotzdem rollte, war ihre Aufmerksamkeit ungleich höher, als wenn der blaue Ball den roten Ball tatsächlich angestoßen hatte.

Die spannende Frage dabei ist: In welchem Alter zeigen Kinder erstmals Anzeichen dafür, dass sie über kausale Erkenntnisse verfügen? Wie die Ergebnisse der vielen systematischen Studien zur Wahrnehmung von Kausalität durchweg belegen, nehmen Kleinkinder bereits im Alter von sechs bis acht Monaten kausale Zusammenhänge wahr (weiterführende Literatur hierzu, siehe bei Muentener & Carey 2010). Allerdings sind bestimmte Aspekte der Kausalität in diesem Alter noch nicht vorhanden. Zum Beispiel scheinen Kleinkinder im Vergleich zu Erwachsenen den Unterschied zwischen zeitlich verzögerten und räumlich auseinanderliegenden Zusammenhängen nicht zu bemerken. Wechselt ein Bewegungsereignis zwischen diesen beiden hin und her, erwacht ihr Interesse nicht erneut. Diese Fähigkeiten der kausalen Wahrnehmung (und andere) bilden sich in einer späteren kindlichen Entwicklungsphase heraus.

Die Ergebnisse bieten vielleicht auch eine Lösung zur Beilegung des Streits zwischen Hume und Kant. Kausalität ist in der Tat eine Besonderheit physikalischer Ereignisse, und wir führen beobachtete Ereignisse in der Tat auf mögliche Ursachen zurück. Doch die Ergebnisse sorgfältiger Untersuchungen der kindlichen Kognition deuten darauf,

dass es damit weit mehr auf sich hat. Die Systematik, mit der Kinder schon im frühesten Alter kausale und nichtkausale Ereignisse auseinanderhalten können, sowie die verfeinerte Unterscheidung in Subsysteme kausaler Ereignisse, die in späten Kinderjahren erfolgt, legt vielmehr nahe, dass es sich um grundlegendes Wissen und nicht um „Illusionen" handelt.

Was also haben wir über das logisch kausale Denken und Schließen des Menschen gelernt? Wir haben gelernt, dass es sich unter zwei Stichpunkten zusammenfassen lässt. Erstens: Wir reagieren sensibel auf Covariationen zwischen Ereignissen und stützen kausale Schlüsse (Induktionen, vom speziellen Fall auf allgemeinere Voraussetzungen) auf die Stärke der Covariation. Zweitens: Wir brauchen einen plausiblen Zusammenhang, der die zwei Ereignisse miteinander verknüpft. Sind beide Bedingungen erfüllt, gehen wir von einem Fall vollendeter Kausalität aus – auch wenn der plausible Zusammenhang sich nur als scheinbar (nicht echt) oder nicht angemessen validiert erweist.

Anders ausgedrückt: Wir folgern kausale Zusammenhänge (oder nehmen diese wahr), indem wir uns beobachtbare Covariationen zwischen möglichen Ursachenfaktoren und dem Effektereignis zunutze machen. Dies führt unter vielen Bedingungen meist zu falsch-positiven und weniger zu falsch-negativen Beurteilungen des Kausalzusammenhangs. Und das ist auch gut so, denn falsch-negative Beurteilungen können sehr viel verheerendere Auswirkungen auf das menschliche Überleben haben als falsch-positive.

Was ist denn das für ein schwarzer Fleck dort im Gebüsch? Ein lauernder Räuber oder ein einfacher Schatten? Eine falsch-positive Beurteilung führt zu dem Schluss, dass es sich um einen Räuber handelt (wenn es gar kein Räuber

ist), und treibt zur Flucht. Eine falsch-negative Beurteilung führt zu dem Schluss, dass es bloß ein harmloser Schatten ist (wenn es in Wahrheit ein Räuber ist), und treibt nicht zur Flucht. Es ist also besser, auf Nummer sicher zu gehen. Ihnen ist schlecht? Wieso? Liegt es vielleicht an dem Essen, das irgendwie komisch geschmeckt hat oder ist es reiner Zufall? Eine falsch-positive Beurteilung führt zu dem Schluss, dass es am komischen Essen liegt (wenn es in Wahrheit gar nicht daran liegt), und Sie es fortan meiden werden. Eine falsch-negative Beurteilung führt zu dem Schluss, dass es reiner Zufall ist (wenn es in Wahrheit doch am Essen liegt), und Sie weiterhin davon essen.

In puncto Überleben zahlen sich falsch-positive Beurteilungen also allemal mehr aus als falsch-negative. Und auch *in puncto* Kausalität lohnt es allemal, wenn wir uns an verlässlichen covarianten Faktoren orientieren, die als Ursache für ein Ereignis relevant sein können („Von diesem Essen ist mir schlecht"), anstatt sie wertneutral zu verwerfen („Ist wahrscheinlich nur Zufall, dass mir jetzt schlecht ist").

Das Problem dabei: Unsere Wissensbasis soll mit *wahren* Überzeugungen gefüllt werden, insbesondere mit *wahren* Überzeugungen auf die Frage „Was verursacht was". Und dies bedeutet, dass wir den Wahrheitsgehalt unserer Überzeugungen überprüfen müssen. Wie geschieht das? Mit einem ganz bestimmten Verfahren, dem sogenannten Hypothesentest: Annahme einer wahren Überzeugung, Ableiten einer Prognose (Deduktion), Überprüfen der abgeleiteten Prognose. Und das ist nicht gerade unsere Stärke, wie wir gleich noch sehen werden. Wenn Sie wissen wollen, wie sich der Wahrheitsgehalt von Hypothesen (oder jedweden anderen Aussagen) überprüfen lässt, so ist das nächste Kapitel für Sie das wohl wichtigste im ganzen Buch.

7

Hypothesentests

Wahrheit und Beweis

1960 berichtete der britische Forschungspsychologe Peter Wason von einem interessanten Phänomen im menschlichen Denken. Er legte seinen Versuchspersonen diese einfache Aufgabe vor:

> Ihnen werden drei Zahlen vorgelegt, die einer bestimmten Regel entsprechen, die ich mir ausgedacht habe. Diese Regel beschreibt eine Beziehung zwischen den drei Zahlen und sie bezieht sich nicht auf ihre absolute Größe, das heißt es ist keine Regel in der Art, dass alle Zahlen größer (oder kleiner) sind als 50 oder dergleichen.
> Ihre Aufgabe ist es, die Regel herauszufinden, nach der die Folge gebildet wurde. Hierzu sollen Sie Zahlentripel bilden, die Ihrer Meinung nach der Regel entsprechen. Nach jeder Folge, die Sie niedergeschrieben haben, werden ich Ihnen eine Rückmeldung darüber geben, ob Ihre produzierte Folge der gesuchten Regel entspricht oder nicht. Sie dürfen dieses Ergebnis auf dem beigelegten Protokollblatt festhalten. Eine zeitliche Begrenzung gibt es nicht, jedoch sollten Sie versuchen, die Regel mit der Minimalmenge von Zahlenfolgen zu ermitteln.

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_7,

> Denken Sie daran, dass es nicht darum geht, Zahlen zu finden, die der Regel entsprechen, sondern dass Sie die Regel dahinter herausfinden sollen.
> Geben Sie die gesuchte Regel erst dann als Ergebnis bekannt, wenn Sie aufgrund Ihrer Versuche ziemlich sicher sind, sie gefunden zu haben, und nicht vorher.
> Hier die Zahlen, die der Regel entsprechen, die ich mir ausgedacht habe: 2 4 6

Bevor Sie nun weiterlesen, raten Sie doch einmal, wie die Regel lautet. Schreiben Sie einige Zahlentripel auf, die Ihnen helfen könnten herauszufinden, ob Sie richtig liegen oder nicht. Will heißen: Würden Sie diese Zahlentripel Wason vorlegen, könnte er Ihnen für jedes einzelne sagen, ob es ein Beleg ist für die Regel, die er sich ausgedacht hatte.

Bestätigungsfehler: Sag, dass ich Recht habe!

Wenn Sie zu der Mehrheit von Wasons Versuchspersonen gehören (54 %), dann lautet die Regel oder die Hypothese, die Sie vorschlagen, vielleicht so oder so ähnlich: „aufsteigende Zahlenfolge im Intervall 2“ oder „Folge von geraden Zahlen“. Doch damit liegen Sie falsch.

Hinzu kommt, dass Sie, wie die meisten von Wasons Versuchspersonen auch, eine bestimmte Hypothese gebildet haben, und dann in der Mehrzahl der Zahlentripel, die Sie formuliert haben (nämlich 77 %), fast ausschließlich *positive Instanzen* für diese Hypothese produziert haben, Instanzen also, die Ihre Hypothese bestätigen, da sie eine oder

mehr Eigenschaften aufweisen, die durch Ihre Hypothese beschrieben ist. Sagen wir mal, Sie haben zum Beispiel folgende Zahlenreihen formuliert: „4 6 8“, „22 24 26“ oder „340 342 344“. All diese Instanzen passen auf die Regel „aufsteigende Zahlenfolge im Intervall 2“ oder „Folge von geraden Zahlen“.

Glückwunsch! Damit haben Sie gerade zwei äußerst hartnäckige Vorlieben des rationalen Denkens demonstriert: Sie haben fleißig nach Beweisen gesucht, die Ihre aufgestellte Hypothese *bestätigen*, um sie stetig aufs Neue zu überprüfen, indem Sie in einem fort positive Instanzen generierten (immer mehr gerade Zahlenfolgen), anstatt Ihre Strategie umzukehren und nach Beweisen zu suchen, die Ihre aufgestellte Hypothese widerlegen, sprich, die sie nicht bestätigen (Folgen von ungeraden Zahlen beispielsweise). Aber ist Ihre Hypothese korrekt? Und war Ihre beweissuchende Strategie optimal?

Wason hätte Ihnen tatsächlich geantwortet, dass jedes Zahlentripel, das Sie generiert haben, ein Beispiel sei für die Regel, die er sich ausgedacht hatte. Doch leider ist die Regel, die er sich ausgedacht hatte, nicht die Regel, die Sie vorgeschlagen haben. Also ist Ihre Hypothese – Ihre geglaubte Meinung – falsch.

Nehmen wir stattdessen an, Sie hätten versucht, Ihre Hypothese zu *falsifizieren*. Dann hätten Sie nach Instanzen gesucht, die gegen die Regel, die Sie im Sinn hatten, eindeutig verstoßen, wie beispielsweise „1 3 8“. Sie hätten also sogenannte *negative Instanzen* generiert, die Eigenschaften aufweisen, die nicht auf Ihre hypothetische Regel passen (Folgen von ungeraden Zahlen beispielsweise).

Und Sie wären wohl überrascht gewesen zu hören, dass „1 3 8“ tatsächlich der Regel entspricht, die Wason sich ausgedacht hatte. Diese Erkenntnis ist sehr nützlich. Sie zeigt nämlich, dass die Zahlentripel überhaupt nicht aus aufeinanderfolgenden geraden Zahlen bestehen müssen.

Schauen wir uns ein weiteres (falsifizierendes) Zahlentripel an, das gegen den letzten kleinen Rest Ihrer geglaubten Meinung verstößt – und zwar, dass die Zahlenfolge aufsteigend sein muss. Nehmen wir „6 3 1“. Wason würde Ihnen antworten, dass dieses Zahlentripel keine Instanz für die Regel ist, die er sich ausgedacht hatte. Diese Antwort ist ebenfalls sehr nützlich. Sie deutet nämlich darauf, dass die Reihe aufsteigend sein muss. Und damit lautet die reine und einfache Regel: *jegliche aufsteigende Zahlenreihe*.

Diese simple Demonstration zeigt etliche Fallen, in die wir tappen, wenn wir Hypothesen (den Wahrheitswert von Aussagen) überprüfen und dabei unserer natürlichen Intuition folgen. Erstens wird klar, dass die ersten Fallreihen, die uns begegnen, unsere Hypothesen verzerren können, da wir sie allzu stark (oder in Einzelfällen auch zu schwach) gewichten. Zweitens wird klar, dass wir, sobald wir eine Annahme oder Hypothese entwickelt haben, nach Beweisen suchen, die uns bestätigen, dass wir richtig liegen, anstatt nach Beweisen zu suchen, die belegen, dass wir damit auch falsch liegen könnten. Dies tun wir, indem wir *positive Instanzen* generieren, um unsere Hypothese zu überprüfen. Und drittens haben diese Studien gezeigt, dass wir selbst angesichts gegenteiliger Beweise offenbar resistent sind, eine geglaubte Annahme aufzugeben und zu ändern: Selbst nachdem sie die Antwort gehört hatten: „Nein, dieses Zah-

lentripel entspricht nicht meiner Regel", hielten viele von Wasons Versuchsteilnehmern an ihrer Hypothese fest und generierten ein weiteres Zahlentripel, um ihre Hypothese mit einem neuen Beispiel verifiziert zu sehen.

Realitätsnahe Studien zum Nachweis von Bestätigungsfehlern

Vielleicht erscheint Ihnen diese Studie etwas gekünstelt. Ist doch egal, wie wir den Wahrheitsgehalt willkürlicher mathematischer Regeln überprüfen, die bloß dem Hirn eines Forschungspsychologen entsprungen sind, oder nicht? Aber betrachten wir die folgende Studie von Lord et al. (1979).

An dieser Studie nahmen 48 Universitätsstudenten teil. Von ihnen waren 24 Befürworter der Todesstrafe, die von deren Abschreckungswirkung überzeugt waren und diese Überzeugung durch einen Großteil der relevanten Studien unterstützt sahen. Die anderen 24 waren Gegner der Todesstrafe, bezweifelten deren Abschreckungswirkung und sahen ihre Überzeugung durch relevante Studien unterstützt. Den Probanden wurde eine kurze Aussage einer Studie über die Todesstrafe präsentiert, wie etwa die folgende:

> Kroner und Phillips (1977) verglichen die Mordraten in 14 US-Bundesstaaten miteinander, und zwar im Jahr vor und im Jahr nach Einführung der Todesstrafe. In elf von 14 dieser Staaten fiel die Mordrate nach Einführung der Todesstrafe niedriger aus. Diese Studie stützt die Abschreckungswirkung der Todesstrafe.

Den Teilnehmern wurde mitgeteilt, die Studie sei empirisch belegt, obwohl sie tatsächlich fiktional war. Anschließend wurden sie gebeten, die beiden folgenden Fragen zu beantworten, und zwar unter Verwendung einer Bewertungsskala von –8 (noch stärker dagegen) bis +8 (noch stärker dafür):

> Hat diese Studie Ihre Einstellung gegenüber der Todesstrafe verändert?
> Hat diese Studie Ihre Überzeugung hinsichtlich der Abschreckungswirkung der Todesstrafe beeinflusst?

Im Anschluss daran erhielten die Teilnehmer nähere Informationen zu dieser Studie wie etwa zur Vorgehensweise, zu Wiederholungsergebnissen oder auch Kurzversionen scharfer Kritiken sowie die Entkräftung derselben durch die Urheber der Studie. Danach sollten sie anhand einer Skala von –8 (sehr schlecht) bis +8 (sehr gut) beurteilen, wie gut oder schlecht sie die Durchführung der Studie fanden, wie überzeugend die Studie in Bezug auf die Abschreckungswirkung der Todesstrafe ihrer Meinung nach war (von –8 für absolut nicht überzeugend bis +8 für absolut überzeugend) und abschließend schriftlich erklären, warum sie glaubten, dass die Studie die Wirksamkeit der Todesstrafe auf die Kriminalitätsrate stützte oder nicht.

Danach wurde der Versuch noch einmal durchgeführt, und zwar mit einer zweiten fiktionalen Studie, die eine gegenteilige Aussage hatte:

> Palmer und Crandall (1977) verglichen die Mordraten von zehn benachbarten US-Bundesstaaten paarweise, immer ein Staat mit und ein Staat ohne Todesstrafe. In acht von

zehn Paaren war die Mordrate in dem Bundesstaat mit Todesstrafe höher. Diese Studie widerlegt die Abschreckungswirkung der Todesstrafe.

Anschließend wurde jeweils der Hälfte der Befürworter wie auch Gegner als erstes eine (fiktionale) Studie „pro" Todesstrafe präsentiert, der anderen Hälfte der Teilnehmer jeweils eine Studie „contra" Todesstrafe.

Dieser Versuchsaufbau kommt einer klassischen Erörterung gleich, bei der ein eigener Standpunkt zu einer Fragestellung gefunden und argumentativ dargelegt werden soll, und dürfte den studentischen Probanden insofern vertrauter gewesen sein. In der Realität verfahren politische Entscheidungsträger und Gesetzesmacher in ihren Debatten um wichtige Themen genau so (auch sie entwerfen Szenarien, deren Inhalte dann in konkrete Entscheidungen einfließen). Doch wie fiel das Ergebnis dieses Versuchs nun aus?

Das bemerkenswerte Ergebnis lässt sich in einem Satz zusammenfassen: Beide Gruppen, sowohl Befürworter als auch Gegner der Todesstrafe, fanden diejenigen Studien überzeugender und beweiskräftiger, die ihre eigene Meinung zur Todesstrafe stützen und belegen konnten, und hatten sich in ihrer anfänglichen Einstellung noch stärker polarisiert.

Die Forscher schlossen daraus, dass wir empirische Befunde relevant verzerrt in Richtung der eigenen Meinung wahrnehmen (*confirmation bias*). Diese Verzerrung zeigt sich in einer auffälligen Neigung, bestätigende Informationen verstärkt wahrzunehmen (*Bestätigungsfehler*), während nichtbestätigende Informationen einer kritisch bezweifelnden Beurteilung unterzogen werden. Und so führen selbst

objektive und wissenschaftlich belegte Aussagen nicht etwa dazu, dass wir die eigene, tief verankerte Meinung infrage stellen, sondern wir polarisieren sie umso extremer. Oder, um es mit den Worten von Sir Francis Bacon zu sagen (1960):

> „Der menschliche Verstand zieht in das, was er einmal als wahr angenommen hat, weil es von Alters her gilt und geglaubt wird, oder weil es gefällt, auch alles Andere hinein, um Jenes zu stützen und mit ihm übereinstimmend zu machen. Und wenn auch die Bedeutung und Anzahl der entgegengesetzten Fälle grösser ist, so bemerkt oder beachtet der Geist sie nicht oder beseitigt und verwirft sie mittelst Unterscheidungen zu seinem grossen Schaden und Verderben, nur damit das Ansehn jener alten fehlerhaften Verbindungen aufrecht erhalten bleibe.."[1]

Box 7.1 Wir sehen, was wir sehen wollen

Wer nun noch immer nicht so richtig glauben will, dass unsere Überzeugungen einen starken Einfluss nehmen auf die Folgerungen, die wir aus objektiven Fakten ziehen, der betrachte am besten das Beispiel vom „fliegenden Pferd" (Olsen 2004). Bildliche Darstellungen von galoppierenden Pferden zeigen – aus prähistorischen Zeiten bis in die Mitte des 19. Jahrhunderts hinein – durchweg Pferde mit gestreckten Beinen – die Vorderbeine greifen weit nach vorne, die Hinterbeine strecken sich weit nach hinten. Und die Menschen wussten sogleich, dass das Bild ein Pferd im Galopp zeigt, denn genau *so* sahen sie ein galoppierendes Pferd. Die Höhlenmenschen sahen es so, Aristoteles sah es

[1] Bacon F (2013) Große Erneuerungen der Wissenschaften. Erstes Buch, S. 40

so und auch der Adel der viktorianischen Zeit. Bis, ja bis es Eadweard Muybridge 1878 gelang, mithilfe von zwölf nebeneinander aufgestellten und durch gespannte Zugdrähte nacheinander auslösenden Fotoapparaten in weniger als einer halben Sekunde zwölf Bilder eines galoppierenden Pferdes aufzunehmen und die Bewegungsabläufe genauestens sichtbar zu machen. Muybridges Fotos zeigten zweifelsfrei, dass sich bei einem galoppierenden Pferd alle vier Beine kurzzeitig in der Luft befinden. Doch im Gegensatz zu den vielen bildlichen Darstellungen von einst ist dies nicht im Moment der weit ausgreifenden und gestreckten Beinen der Fall, sondern dann, wenn sich alle vier Beine unter dem Pferdekörper befinden. Man spricht vom „Moment des schwebenden Übergangs". Auch heute noch zeichnen Kinder galoppierende Pferde so wie man es vor Muybridges Aufnahmen tat.

Wenn das Gehirn die Wahrnehmung verzerrt

Kurz vor der Präsidentschaftswahl 2004 führte eine Gruppe von Forschern eine MRT-Studie mit 30 männlichen Probanden durch, von denen sich die eine Hälfte als überzeugte Republikaner und die andere Hälfte als überzeugte Demokraten beschrieben (Westin et al. 2006). Während das MRT Bilder ihrer Hirnaktivität aufzeichnete, wurden sie gebeten, widersprüchliche Äußerungen sowohl von George W. Bush als auch von John Kerry zu bewerten. Ihnen wurden nacheinander 18 präparierte Aussagen (Stimuli) präsentiert, von denen sich jeweils sechs auf Präsident George W. Bush bezogen, sechs auf seinen Herausforderer Senator John Kerry und weitere sechs auf eine politisch

neutrale männliche Kontrollperson (wie beispielsweise auf den Schauspieler Tom Hanks). Jede Versuchsreihe begann damit, dass die Männer zunächst nur eine Aussage von einer dieser drei Personen zu lesen bekamen. Auf diese Aussage folgte eine evidenzbasierte Information darüber, dass die betreffende Person Dinge getan habe, die mit ihrer ursprünglichen Aussage nicht vereinbar seien. Nach Erhalt dieser Information gingen die meisten Probanden davon aus, dass die betreffende Person unehrlich sei und Wählerstimmen fischen wolle nach dem Motto: „Das eine sagen, das andere tun".

In einem nächsten Schritt sollten die Männer bewerten, in welchem Maß Worte und Taten ihrer Meinung nach in Widerspruch standen. Zum Schluss wurde ihnen eine andere Aussage präsentiert, die den geglaubten Widerspruch aufhob, und sie wurden erneut gebeten zu bewerten, in welchem Maße sich Worte und Taten ihrer Meinung nach widersprechen.

Die mündlich gegebenen Antworten der Probanden zeigten ein Muster emotionsbedingter Verzerrungen im rationalen Denken (*emotional bias*): In Bezug auf ihren favorisierten Kandidaten leugneten sie, offensichtliche Widersprüche bemerkt zu haben, die ihnen beim Gegenkandidaten jedoch sofort aufgefallen waren. In Bezug auf eine neutrale Kontrollperson (wie Tom Hanks) hingegen unterschieden sich beide Lager in ihren Reaktionen auf widersprüchliche Aussagen nicht.

Dass Wahrnehmung und mithin (rationales) Denken und Urteilen der Probanden derart verzerrt waren, liegt schlicht daran, dass diese Vorgänge eher von emotionalen denn von kognitiven Prozessen gesteuert werden: Die

MRT-Bilder zeigten, dass der Teil des Gehirns, der für das logische Denken zuständig ist, der dorsolaterale präfrontale Cortex (dlPFC), bei der Bewertung der vorgelegten Aussagen *nicht* involviert war. Stattdessen zeigten sich diejenigen Hirnareale am aktivsten, die an der Verarbeitung von *Emotionen* (der ventromediale präfrontale Cortex, VMPC), an *Konfliktlösungen* (der anteriore cinguläre Cortex, ACC) und an *moralischen Urteilsbildungen* (der posteriore cinguläre Cortex, PCC) beteiligt sind. Als die Probanden schließlich zu ihrem letztgültigen, wenngleich völlig verzerrten Urteil gelangten (indem sie Informationen, die ihren vorgefassten Überzeugungen widersprachen, eisern ignorierten), verringerte sich die Aktivität jener neuronalen Schaltkreise, die negative Emotionen (wie Betrübnis oder Abscheu) hervorrufen, wohingegen jene des Belohnungssystems, die für Wohlgefühle verantwortlich sind, verstärkt aktiv waren. Der Leiter dieser Studie fasste die Ergebnisse folgendermaßen zusammen:

> Wir konnten keinerlei erhöhte Aktivität in den Hirnregionen beobachten, die beim logischen Denken normalerweise beteiligt sind. Stattdessen sahen wir ein Netzwerk gefühlsbezogener Schaltkreise aufleuchten, darunter Schaltkreise, die mutmaßlich an der Regulierung von Emotionen beteiligt sind, sowie Schaltkreise, die an der Lösung von Konflikten mitwirken. Keiner der Schaltkreise, die an bewussten rationalen Denkprozessen beteiligt sind, waren sonderlich aktiv. Es schien vielmehr so, als wirbelten unbeugsame Partisanen das ganze kognitive Kaleidoskop durcheinander, bis sie die Ergebnisse bekamen, die sie wollten, die sie dann mit aller Macht zu bekräftigen und zu verstärken suchten,

während sie gleichzeitig negative emotionale Zustände verhinderten und positive Emotionen aktivierten.

Ergebnisse wie diese lassen in aller Deutlichkeit erkennen, dass wir zu einer Verzerrung der Wahrnehmung und damit zu Bestätigungsfehlern tendieren, wenn wir die Wahrheit einer Aussage durch Hypothesentests zu überprüfen suchen. Um diese inhärente Verzerrung zu überwinden, brauchen wir eine Untersuchungsmethode, die diese Tendenz ausgleicht, in etwa so, wie Gegengewichte zum Ausgleich schwerer Ladungen verwendet werden. Die Lösung für dieses Dilemma ist die hypothetisch-deduktive Methode, eine Form der wissenschaftlichen Untersuchung. Sie ist das Ergebnis einiger der schönsten und klarsten Gedanken, die Philosophen und Wissenschaftler im Laufe der Jahrhunderte hervorgebracht haben. Ausgehend von der Tatsache, dass wir von Natur aus voreingenommen sind, wenn es darum geht, die eigenen Annahmen und Überzeugungen zu prüfen, wurde eine wissenschaftliche Methode geboren, die jegliche Überzeugungen (im wissenschaftlichen Sinne *Hypothesen*) auf einen rationalen Prüfstand stellt, um sie auf ihre innere Widerspruchsfreiheit hin zu überprüfen. Dies geschieht, indem man wiederholbare Experimente durchführt und damit nach Evidenzen sucht, die der zu beweisenden Hypothese *widersprechen* (die sie im wissenschaftlichen Sinne *falsifizieren*). Es lohnt sich daher durchaus, diese schwierige Geburt in der Wissenschaftsgeschichte ein wenig näher zu beleuchten, um sich über die Neigung, die eigenen Ergebnisse subjektiv in die Richtung zu verzerren, in der man sie haben will (*observer bias*), etwas klarer zu werden.

Der (historische) Weg der Wissenschaft

Forschergeist will Wahrheit wissen – oder, um es mit dem schönen Satz von Charles Peirce, dem Vater der wissenschaftlichen Philosophie in den Vereinigten Staaten, zu sagen: „Radikales Fragen kennt unbedingte Voraussetzungen nur als Denkmittel“ (1877).[2] Das Fragen beginnt aus einem Zustand der Ungewissheit heraus und bewegt sich hin zu einem Zustand der Gewissheit. Wir beenden das Fragen erst, wenn wir befriedigt wissen, was wir zu wissen verlangten – die Wahrheit.

Es gibt viele Mittel und Wege, die Ungewissheiten unserer Vorstellungen und Annahmen zu verringern, um „gesicherte“ Erkenntnisse zu gewinnen. Wir könnten es beispielsweise mit Platon halten (429–347 v. Chr.) und auf geistige Reflektion und Vernunft setzen, die einzigen beiden (Denk-)Mittel, durch die nach seiner Überzeugung die wahre Natur aller Dinge erkennbar sei. Wir könnten aber auch seinem berühmtesten Schüler Aristoteles (348–322 v. Chr,) folgen und die Idee hochhalten, dass Wahrheit durch Erfahrung gefunden werden könne, insbesondere durch genaue Beobachtungen und Messungen. Oder wir gehen noch weiter und schließen uns dem muslimischen Wissenschaftler Alhazen an (965–1040 n. Chr.), der mit zahlreichen kontrollierten Experimenten einen ganz praktischen Ansatz verfolgte, um der Wahrheit auf den Grund zu gehen. Wenn wir Aristoteles und Alhazen folgen, folgen wir dem Pfad *wissenschaftlicher Methoden.*

[2] http://www.gleichsatz.de/b-u-t/begin/cspeirce.html (Zugriff 29.1.2014)

Kernelement der wissenschaftlichen Methodik sind *Hypothesentests*. Man verfährt dabei, indem man aus einer vorgefassten und als wahr vorausgesetzten Annahme (*Hypothese*) eine Vorhersage über Ereignisse formuliert, die in ihrem Verlauf direkt beobachtet werden können. Diese Beobachtungen liefern Beweise (*Evidenzen*), wonach unsere nur angenommene und unbewiesene Hypothese entweder wahr ist oder falsch. Dass wissenschaftliche Sätze nur dann als solche zu betrachten seien, wenn sie an der Erfahrung *verifiziert* (bestätigt) werden können, ist nach Sir Francis Bacon (1561–1626) das ausschlaggebende Kriterium, das diese Methode von anderen wissenschaftlichen Untersuchungsmethoden unterscheidet. Laut Bacon „beweist" man den Wahrheitsgehalt einer bestimmten Hypothese, indem man Beobachtungen sammelt, die mit dieser Hypothese *konform* gehen. Wiederholte Bestätigungen führen zu „Gesetzen" – Bestätigungen der Regeln, die die natürliche Welt „beherrschen". Bacon führte auch das Formulieren von sogenannten Alternativhypothesen ein und führte explizite Untersuchungen durch, um theoretische Fragestellungen voneinander zu unterscheiden. Er war überzeugt, diese Methode bringe der wissenschaftlichen Erkenntnis den größten Fortschritt. Und so basierte seine Methode in erster Linie auf dem Sammeln von konformierenden Beweisen für die aufgestellten Hypothesen.

Angesichts der Studien, die in diesem Kapitel eingangs beschrieben wurden, mögen Sie Bacons Position mit einer gewissen Skepsis begegnen. Da sind Sie nicht allein. David Hume (1711–1776) zieht ebenfalls in Zweifel, dass eine solche Bestätigung überhaupt möglich sei. Er vertritt vielmehr die Meinung, dass „konsequent empiristische Be-

obachtungen“ nicht als „Beweis“ für die Wahrheit einer Hypothese gelten könnten. Um seine Position zu stützen, entwickelte er das heute noch berühmte Argument vom „schwarzen Schwan“: Keine noch so große Menge an induktiven Schlüssen, die wir ziehen, könne je beweisen, dass die Aussage „Alle Schwäne sind weiß“ wahr sei, denn es könne irgendwo immer eine bis jetzt unentdeckte Population schwarzer Schwäne geben. All die Beweise, die gesammelt werden, um den Wahrheitsgehalt dieser Hypothese zu belegen, beweisen lediglich, dass diese Hypothese bisherig gestützt war. (Und wie sich herausstellte, gibt es schwarze Schwäne wirklich – in Australien!)

Karl Popper (1902–1994) verschärft diesen Einwand noch, indem er die empirische Überprüfung als Grundlage von Hypothesentests völlig demontiert. Er stellt den Gegenbeweis über den Beweis und die Falsifikation über die Verifikation. Wahrheit, so argumentiert er, könne nur herausgefunden werden, indem man herausfände, was Wahrheit nicht sei. Keine noch so große Beispielmenge, so Popper, könne die absolute Wahrheit einer Behauptung beweisen, jedoch wachse die relative Wahrheit dieser Behauptung an (d. h. die feste Überzeugung, die sie erweckt), so wie sich die Beispiele gescheiterter Falsifikation zu häufen beginnen würden. Anders ausgedrückt: Je öfter man scheitert, eine Hypothese zu widerlegen, umso stärker wächst das Vertrauen in den Wahrheitsgehalt dieser Hypothese. Manchmal reicht ein einziges Experiment aus, um eine Hypothese zu widerlegen. Ein einziger schwarzer Schwan falsifiziert die Überzeugung, dass alle Schwäne weiß sind. Oder wie Einstein einst sagte: „Keine noch so große Zahl von Experi-

menten kann beweisen, daß ich recht habe; ein einziges Experiment kann beweisen, daß ich unrecht habe."[3]

Beweise mir, dass ich mich irre

Wenn die Falsifikation nun eine so große Bedeutung hat, stellt sich die Frage, warum Hypothesentests in aller Regel so ausgerichtet sind, dass sie nach verifizierenden Beweisen suchen. Vielleicht liegt die Antwort darin, dass sich die methodische Anleitung zur Gewinnung neuer Erkenntnisse (Heuristik), auf der solche Hypothesentests basieren, besonders effizient und effektiv nutzen lässt, auch wenn wir am obigen Beispiel der 2-4-6-Aufgabe gesehen haben, dass sie zu systematischen Fehlern führen kann.

Die weite Verbreitung und die Folgen dieser Vorgehensweise sind sehr schön in einem Aufsatz von Klayman und Young-Won Ha (1987) beschrieben. Betrachten wir noch einmal Wasons Zahlenfolge 2 4 6. Wason hatte eine Regel im Kopf, die eine Folge von Zahlen exakt beschreibt, die wir die Zielreihe (*target set*, TS) nennen wollen. Seine Regel, die diese Zielreihe beschreibt, nennen wir die Zielregel (*target rule*, TR), die da lautete: „jegliche aufsteigende Zahlenreihe". Dies war die Regel, die Sie herausfinden sollten, indem Sie Tests generierten. Um die Zielregel zu finden, konstruierten Sie Hypothesen, also Vermutungen darüber, wie es sein könnte, und basierten Ihre sicherste Vermutung auf dem ersten Zahlentripel (2 4 6), das Sie zu sehen bekommen hatten. Ihre sicherste Vermutung ist eine

[3] http://gutezitate.com/zitat/230485 (Zugriff 29.1.2014)

hypothetische Regel (*hypothesized rule*, HR). Ihre hypothetische Regel beschreibt eine Reihe von Zahlen ebenfalls ganz exakt. Wenn Ihre hypothetische Regel „aufsteigende gerade Zahlen" lautete, dann generierten Sie hypothetische Zahlenreihen (*hypothesized set*, HS), die Ihre hypothetische Regel erfüllten.

Und so verfahren Sie sehr zielgerichtet: Ihr Ziel ist, Ihre hypothetische Regel in Übereinstimmung zu bringen mit der gesuchten Zielregel (HR = TR). Das würde bedeuten, dass jede hypothetische Zahlenreihe exakt auf die Zielreihe passt (TS = HS). Sie könnten sodann nach einer von zwei Strategien für Hypothesentests vorgehen. Sie könnten ein Zahlentripel vorschlagen, von dem sie glauben, dass es die Eigenschaften der Zielreihe hat (z. B. 6 8 10). Diese Strategie nennt man positive Teststrategie (Verifikation). Oder Sie könnten ein Zahlentripel vorschlagen, von dem Sie glauben, dass es *nicht* die Eigenschaften der Zielreihe hat (z. B. 2 4 7). Diese Strategie nennt man negative Teststrategie (Falsifikation). Wason fand heraus, dass die Probanden in der großen Mehrheit nach der positiven Teststrategie verfuhren. Sie überprüften ihre Hypothesen, indem sie Testreihen generierten, die vermutlich die Eigenschaften der Zielreihe aufwiesen (z. B. gerade Zahlen).

Klayman und Ha argumentierten, dass es die Anwendung der positiven Teststrategie unter bestimmten Bedingungen (Bedingungen, die tatsächlich sehr realistisch sind) erlaube, Versuche anzustellen, die die schlüssigsten Beweise für diese Hypothese liefern. Oder anders gesagt: Diese unter bestimmten Bedingungen angestellten Versuche erlauben es Ihnen, den erwarteten Erkenntnisgewinn zu maximieren. Hypothesentests verlangen Anstrengungen, wes-

halb es sehr klug von Ihnen ist, Versuche anzustellen, von denen Sie sich den größtmöglichen Erkenntnisgewinn bei geringstmöglicher Anstrengung erwarten.

Die wichtigste unter diesen „bestimmten Bedingungen" ist die *Seltenheitsbedingung* – das heißt, das, was Sie untersuchen, ist ein seltenes Ereignis. In einem solchen Fall nach falsifizierenden Instanzen zu suchen, gleicht der sprichwörtlichen Suche nach der Nadel im Heuhaufen. Das liegt daran, dass die Menge derjenigen Elemente, welche die Eigenschaften besitzen, die der korrekten Regel entsprechen, sehr klein ist, während die Menge derjenigen Elemente, die jene Eigenschaften nicht besitzen, extrem groß ist. Unter der Bedingung der Seltenheit ist es weitaus produktiver, die Elemente zu untersuchen, die die von Ihnen vermuteten Eigenschaften besitzen.

Betrachten wie den Fall der Depression. Stellen wir uns einmal vor, Sie sind die erste Person, die die Vermutung hegt, dass eine der Ursachen depressiver Erkrankungen ein zu niedriger Serotoninspiegel ist. Ihre vermutete Regel würde in etwa so lauten: „Ein niedriger Serotoninspiegel führt zu einer Depression." Stellen wir uns weiterhin vor, dass Sie die Gelegenheit bekommen, Ihre Hypothese an vier Patienten zu überprüfen, die zu Ihnen in die Klinik kommen.

Der erste Patient weist einen niedrigen Serotoninspiegel auf, der zweite einen normalen, beim dritten wurde bereits eine Depression diagnostiziert und der vierte ist nicht depressiv, sondern kam aus anderen Gründen in die Klinik. Sie haben nicht viel Zeit, weshalb sie rasch entscheiden müssen, welchen oder welche Patienten Sie untersuchen. Wen würden Sie aussuchen?

Diese Variante des Hypothesentest bezeichnet man als das Vier-Karten-Problem oder Wasons Kartenaufgabe (*Wason card selection*), da es von Peter Wason (dem Mann, der uns auch die 2-4-6-Aufgabe stellte) in den frühen 1060er-Jahren erstmals vorgebracht und untersucht wurde. Wenn Sie sich so entscheiden wie der Großteil der Probanden in Wasons Studie (und seither in zahllosen weiteren), wählen Sie den Patienten mit dem niedrigen Serotoninspiegel, um zu überprüfen, ob der Patient depressiv ist, sowie den Patienten, der bereits an einer Depression leidet, um herauszufinden, ob sein Serotoninspiegel ebenfalls niedrig ist. So weit, so gut. Jedoch werden Sie sich wohl den Vorwurf gefallen lassen müssen, dem oben beschriebenen Bestätigungsfehler (*confirmation bias*) aufgesessen zu sein. Schließlich haben Sie damit genau jene Fälle ausgewählt, die Ihre in Betracht gezogene Hypothese bestätigen könnten. Mal angenommen, Sie untersuchen den depressiven Patienten und stellen fest, dass sein Serotoninspiegel tatsächlich niedrig ist. Dieses Ergebnis würde mit Ihrer Hypothese übereinstimmen. Aber was, wenn sein Serotoninspiegel normal ist? Würde das Ihre Hypothese widerlegen? Wohl nicht. Ihre Hypothese besagt nämlich nicht, dass ein niedriger Serotoninspiegel die einzige Ursache für eine Depression ist. Der Patient könnte auch aus anderen Gründen depressiv sein.

Was aber wäre, wenn Sie sich entschieden hätten, den nichtdepressiven Patienten zu untersuchen, und stellten nun fest, dass sein Serotoninspiegel niedrig ist. Sie hätten damit einen unwiderlegbaren Beweis dafür, dass Ihre Hypothese falsch ist. Insofern scheint dieser Test doch sehr viel effizienter – und aussagekräftiger.

Ein Logiker könnte argumentieren, dass Ihre aufgestellte Hypothese einer wahrheitsfunktionalen Bedingung in der Form „wenn P, dann Q" untergeordnet ist, das heißt, dass die Wahrheitsfunktion für eine konditionale Aussage nur dann falsch ist, wenn der Bedingungssatz (P) wahr ist, aber der Folgesatz (Q) falsch. Eine optimale Strategie für einen Hypothesentest wäre daher die, den Patienten mit dem niedrigen Serotoninspiegel zu untersuchen sowie den Serotoninspiegel desjenigen Patienten, der nicht depressiv ist. Würde sich im Ergebnis ein niedriger Serotoninspiegel, aber keine Depression zeigen, wäre dies ein unwiderlegbarer Beweis dafür, dass der Bedingungssatz falsch war. Doch wie wir gesehen haben, verfahren die meisten Probanden genau umgekehrt und versuchen eine aufgestellte Regel (Hypothese) zu bestätigen und nicht zu widerlegen. Mal ehrlich, wären Sie darauf gekommen? Und noch eine Frage: Ist diese Denk- bzw. Vorgehensweise rational im Sinne des Ausmaßes, mit welchem wir imstande sind, aus einer gegebenen Menge von Behauptungen alle Folgen alleine durch Logik zu erkennen, ohne uns dabei von möglichen, aber falschen Schlüssen verführen zu lassen (Anm. d. Übers.)?

Nähern wir uns diesem Thema aus der anderen Perspektive und wenden eine positive Teststrategie an. Diese positive Strategie ist sinnvoll, wenn das zu behandelnde Phänomen oder Ereignis selten auftritt. Was wäre denn, wenn es nicht um einzelne Patienten ginge, sondern um Experimente, die Sie durchführen könnten? Die Anwendung einer negativen Teststrategie (Falsifikation) würde bedeuten, den Serotoninspiegel derjenigen Patienten zu überprüfen, die nicht depressiv sind, um zu versuchen, die aufgestellte Hypothese zu falsifizieren. Depressive Erkrankungen sind

jedoch, relativ gesehen, selten. Die Depressionsrate der Erwachsenen in den USA ab einem Alter von 18 Jahren liegt bei rund 10 %. Demnach sind 90 % der erwachsenen US-Amerikaner nicht depressiv. Bei einer Stichprobe mit 1000 Menschen würde man also erwarten, dass nur rund 100 depressiv und 900 nicht depressiv sind. Es wäre daher effizienter, den Serotoninspiegel der depressiven Menschen zu untersuchen, da diese Menge kleiner ist und sie Menschen mit der relevanten, sprich depressiven Eigenschaft enthält. Würden die Serotoninspiegel in dieser Gruppe ungewöhnlich niedrig ausfallen, würde dies für niedrige Serotoninspiegel bei depressiven Erkrankungen sprechen. Ihre Arbeit ist damit aber nicht getan, denn Sie wollen diese Feststellung natürlich mit weiteren Experimenten untermauern (die wir gleich noch erläutern werden). Doch fürs Erste liefert diese positive Teststrategie mit einer kleinen Menge und vergleichsweise wenig Aufwand ein sehr viel größeres Maß an nützlichen Informationen als eine groß angelegte negative Teststrategie mit einer großen Menge.

Nach Oaksford und Chater (1999) gehen wir nicht zwangsläufig irrational vor, wenn wir unser Augenmerk auf potenziell bestätigende Informationen lenken. Vielmehr ist in diesem Fall, wenn zudem positive Instanzen der Theorie seltener sind als negative, die Suche nach positiven Instanzen der Theorie durchaus rational, weil sie häufiger zur Falsifizierung der Hypothese führt und damit den erwarteten Informationsgewinn maximiert. Dies entspricht aber eher der Theorie der optimalen Datenselektion (ODS) nach der Bayes'schen Entscheidungsregel (s. Kap. 1) und nicht dem von Karl Popper entwickelten Falsifikationismus (kritischer Empirismus).

Oaksford und Chater (1999) schlagen vier Schritte für Hypothesentests nach dem Modell der optimalen Datenselektion vor:

1. Datenziele: Als Ziel wird gesetzt, die Daten zu selektieren, die die größte erwartete Aussagekraft (*greatest expected informativeness*, EI_g) darüber haben, ob die Regel wahr ist (oder ob der Bedingungssatz (P) und der Folgesatz (Q) der Regel voneinander unabhängig sind).
2. Datenumgebung: Man betreibt Hypothesentests normalerweise dann, wenn außergewöhnliche Ereignisse vorliegen. Dies bedeutet, dass man bereits zu Beginn der Hypothesentests davon ausgeht, dass die relevanten Eigenschaften in der Umgebung selten sind (z. B. Depression oder niedrige Serotoninspiegel).
3. Rechnerbegrenzung: Wenn wir unbegrenzt Zeit und Geld hätten, könnten wir alle möglichen Untersuchungen anstellen, um der Wahrheit auf den Grund zu gehen. Doch das haben für gewöhnlich nicht. Stattdessen müssen wir Entscheidungen in Echtzeit mit begrenzten Mitteln treffen.
4. Datenoptimierung: Man selektiert möglichst viele aussagekräftige Informationen. Unter der Annahme der Seltenheit lassen sich für Wasons Vier-Karten-Problem nacheinander folgende Aussagen über den erwarteten Informationsgewinn formalisieren: $EI_g(P) > EI_g(Q) > EI_g(\text{Nicht-Q}) > EI_g(\text{Nicht-P})$. Genau darum schien Ihnen die Notwendigkeit, den Fall P (Depression) und den Fall Q (niedriger Serotoninspiegel) zu untersuchen, so naheliegend.

Stellt sich jedoch im Zuge der Datenselektion die Vermutung der Seltenheit als falsch heraus, so verlangt die ODS, die Datenselektion erneut vorzunehmen. Dieser Ablauf wurde vielfach experimentell beobachtet. Beispiel: Fordert man die Probanden auf, Fälle von Patienten mit Depression und Fälle von Patienten ohne Depression zu prüfen, steigt die Wahrscheinlichkeit, dass die Probanden die nichtdepressiven Patienten zur Prüfung auswählen, mit zunehmender Zahl der depressiven Patienten. Anders formuliert: Die Probanden entscheiden sich immer für die Prüfung der seltenen Fälle, will heißen, wenn die nichtdepressiven Fälle zunehmend selten erscheinen, werden eher geprüft als die zunehmende Zahl der depressiven Fälle (Oaksford et al. 1997).

Gut, ich werde es dir beweisen!

Kommen wir noch einmal zurück auf unsere Hypothese bezüglich des Zusammenhangs von Serotoninspiegel und Depression. Angenommen, Sie haben nun eine Reihe von depressiven Menschen (seltene Fälle) untersucht und festgestellt, dass sie im Großen und Ganzen tendenziell einen niedrigen Serotoninspiegel aufweisen. Ihre Hypothese scheint aufgrund dieser Information also glaubhaft, aber damit geben Sie sich nicht zufrieden. Sie wollen Beweise.

Sie denken noch einmal darüber nach und kommen plötzlich darauf, dass Ihre Hypothese impliziert, dass ein steigender Serotoninspiegel die Depression doch eigentlich lindern müsse. Also entwickeln Sie ein Arzneimittel, das als Antidepressivum wirkt, einen sogenannten selektiven Serotoninwiederaufnahmehemmer (*selective serotonin reuptake inhibi-*

tor, SSRI), der die Wiederaufnahme von Serotonin im Gehirn (genauer aus dem synaptischen Spalt in die Präsynapse) hemmt. Mit der Prüfung dieses Arzneimittels auf seine Wirksamkeit haben Sie nicht nur ein Medikament zur Behandlung von Depressionen gefunden, sondern auch einen Beleg für den Wahrheitsgehalt Ihrer Hypothese, wonach Depression durch einen niedrigen Serotoninspiegel verursacht wird.

Ob Sie es erkennen oder nicht, aber Sie haben soeben begonnen, die moderne hypothetisch-deduktive Methode der wissenschaftlichen Untersuchung anzuwenden, die wie folgt zusammengefasst werden kann:

1. Identifizieren eines Problems
2. Formulieren einer Vermutung (Hypothese), um das Problem zu erklären
3. Folgern oder Ableiten der Vorhersage(n) aus dieser Hypothese
4. Konzipieren eines oder mehrerer Experimente, um jede Vorhersage zu überprüfen

Doch wenn Sie strikt danach verfahren, laufen Sie Gefahr, einmal mehr nur nach bestätigenden Beweisen für Ihre Hypothese zu suchen. Also optimieren Sie das Ganze ein wenig, indem Sie sich auch negative Teststrategien erlauben und Falsifikationsversuche zulassen. Dies tun Sie, indem Sie vorab eine weitere Aussage formulieren, die sogenannte Nullhypothese, wobei sich Nullhypothese und Ihre bevorzugte Hypothese *gegenseitig ausschließen*:

H_0: SSRI hat *keine* Auswirkung auf Depression
H_A: SSRI hat *eine* Auswirkung auf Depression

Beachten Sie, dass die beiden Hypothesen nicht gleichzeitig wahr sein können – sie schließen sich gegenseitig aus. Sie beginnen Ihre Untersuchung, indem Sie die depressiven Patienten in zwei Gruppen einteilen. Der einen Gruppe wird ein SSRI verabreicht, der anderen ein Placebo (ein Scheinarzneimittel ohne Wirkstoff). Danach ermitteln Sie für jede Gruppe den jeweiligen Grad der Depression. Die Placebogruppe ist dabei besonders wichtig, denn Placeboeffekte sind real messbare Effekte: Patienten fühlen sich besser, nur weil sie meinen, ein Medikament eingenommen zu haben. Im Fall einiger Erkrankungen, bei Ulkuserkrankungen etwa, verspüren bis zu 50 % der Placebopatienten eine deutliche Besserung ihres Zustands. Um nun beweiskräftig folgern können, dass Ihr SSRI-Medikament tatsächlich wirksam ist, müssen Sie die Besserungsraten der SSRI-Gruppe mit denen der Placebokontrollgruppe vergleichen.

Beachten Sie, dass Ihre bevorzugte Hypothese H_A so formuliert ist, dass Sie nach Veränderungen der depressiven Zustände suchen, und zwar in beide Richtungen – zum besseren wie zu schlechteren hin. Damit nehmen Sie einen sogenannten zweiseitigen Test vor (*two-tailed test*), denn Sie suchen nach Veränderungen in beide Richtungen. Wenn Sie keine Unterschiede feststellen, können Sie die Nullhypothese nicht zurückweisen und folgern daraus, dass der SSRI unwirksam war. Doch wenn die Werte der SSRI-Gruppe niedriger ausfallen als die der Placebogruppe, bedeutet dies, dass Ihr SSRI-Medikament den Grad der Depression verringert hat. Und daraus könnten Sie nun folgern, dass Ihr SSRI-Medikament die Depression tatsächlich gelindert hat. Auf der anderen Seite bedeutet dies, dass Ihr SSRI-Medikament den Grad der Depression erhöht hat

(will heißen, die Depression verschlimmert hat), wenn die Werte der SSRI-Gruppe höher ausfallen als die der Placebogruppe. Ihre so formulierte Nullhypothese veranlasst Sie also, mittels statistischer Tests in beiden Richtungen nach Unterschieden suchen.

Haben Sie Ihre bevorzugte Hypothese jedoch wie folgt formuliert: „Mein SSRI lindert Depression", dann lautet die gegenseitig ausschließende Hypothese: „Mein SSRI hat keinerlei Auswirkungen auf Depressionen und macht sie auch nicht schlimmer". Damit nehmen Sie einen sogenannten einseitigen Test vor (*one-tailed test*), da Sie *nur* nach Besserungen suchen, die auf Ihr SSRI-Medikament zurückzuführen sind. Wenn Sie diese so formulierte Nullhypothese zurückweisen, stellen Sie anheim, dass die SSRI-Gruppe einen vergleichsweise niedrigeren Depressionsgrad aufweist. Sollte Ihr SSRI-Medikament die Depressionen aber tatsächlich verschlimmern, würde Ihnen genau dies entgehen, da Ihre Nullhypothese alle positiven Ergebnisse unter „keinerlei Auswirkung auf Depressionen" subsumiert. Und so können Sie lediglich sagen: „Die Ergebnisse zeigen, dass mein SSRI-Medikament die Depression nicht lindert. Ich vermag aber nicht zu sagen, ob dies darauf zurückzuführen ist, dass es keinerlei Auswirkungen auf Depressionen hat (SSRI-Gruppe und Placebogruppe weisen einen gleich hohen Depressionsgrad auf), oder darauf, dass es Depressionen tatsächlich verschlimmert (die SSRI-Gruppe weist einen höheren Depressionsgrad auf als die Placebogruppe).

Ihnen liegt nun für jeden Teilnehmer in jeder Gruppe eine Messung für den Grad seiner Depression vor. Und weil jeder Mensch ein Original und keine Kopie eines anderen ist, können die einzelnen Messungen innerhalb der Grup-

pe von Mensch zu Mensch auch unterschiedlich ausfallen. Wenn wir die Körpertemperatur von einer Million Menschen messen, stellen wir fest, dass sie nur geringfügig variiert. Manche haben eine Körpertemperatur von 36,8 °C, andere von 37,1 °C und so weiter. Wenn wir diese Menschen nun per Zufallsauswahl in Gruppen einteilen und die jeweilige durchschnittliche Körpertemperatur der Gruppen miteinander vergleichen würden, so würden wir feststellen, dass auch diese Werte unterschiedlich ausfallen. Für die eine Gruppe liegt die Durchschnittstemperatur vielleicht bei 36,8 °C, für die andere bei 37,1 °C. Doch wenn wir den Durchschnitt aller Gruppendurchschnitte nehmen, würden wir feststellen, dass die am häufigsten vorkommende Temperatur bei 37 °C liegt und damit bei der wahren durchschnittlichen normalen Körpertemperatur des Menschen. Mit anderen Worten: Beliebig viele Durchschnitte verteilen sich (tendenziell) um den wahren Mittelwert einer Grundgesamtheit (Population). Diese (Wahrscheinlichkeitsgesetzen folgende) Verteilung von Mittelwerten nennt man auch Stichprobenverteilung (*sample distribution*). Des Weiteren stellen wir für die Mittelwerte unserer Zufallsauswahl fest: Je größer die Gruppe, desto kleiner der Abstand zum Durchschnittswert der Grundgesamtheit; und je kleiner die Gruppe, desto größer die Ausreißer nach oben oder unten.

Stellt man die Verteilung durchschnittlicher Werte (in unserem Fall der Temperatur) grafisch dar, ergibt sich eine glockenförmige Kurve (Glockenkurve), was uns ermöglicht, etwas mathematisch sehr Interessantes und Nützliches zu tun. Wir können diese Verteilung nutzen, um die Wahrscheinlichkeit, falsche Schlüsse zu ziehen, zu begrenzen. Die Glockenform tritt auf, weil sich die Werte um den

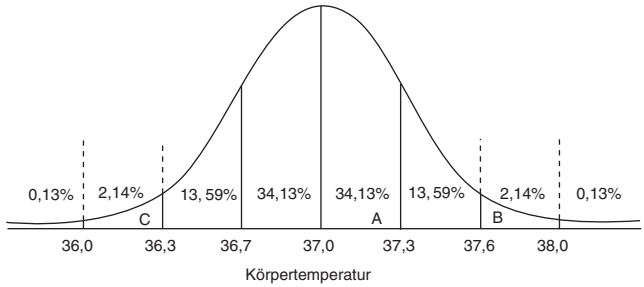


Abb. 7.1 Die Glockenkurve von normalverteilten Werten.

wahren Mittelwert einer Grundgesamtheit herum verteilen, wie in Abb. 7.1 dargestellt.

Die sogenannte Standardabweichung beschreibt das Maß für die typisch auftretende Streubreite um einen Mittelwert. Rund 34 % der Werte liegen zwischen Mittelwert und der einfachen Standardabweichung oberhalb des Mittelwerts, und rund 34 % der Werte liegen zwischen Mittelwert und der einfachen Standardabweichung unterhalb des Mittelwerts. Weitere 14 % liegen zwischen der ein- und zweifachen Standardabweichung oberhalb des Mittelwerts, und 14 % zwischen der ein- und zweifachen Standardabweichung unterhalb des Mittelwerts. Etwa 2 % der Werte liegen zwischen der zwei- und der dreifachen Standardabweichung oberhalb des Mittelwerts; gleiches gilt für Werte unterhalb des Mittelwerts. Insgesamt sind 95 % aller Werte in einem Intervall der Abweichung zwischen +2 und –2 vom Mittelwert zu finden.

Wie stellt sich dies nun dar, wenn wir die Hypothese überprüfen, die besagt, dass eine Person Fieber hat? Als Standardabweichung nehmen wir den Wert 0,6, was ungefähr auch stimmt. Wir haben drei Personen – A, B und C –,

deren Körpertemperatur in der Glockenkurve von Abb. 7.1 zu sehen ist.

Da die 37,2 °C von A nicht so arg weit entfernt sind von 37 °C, liegt der Wert noch gut innerhalb der Streubreite der normalen Temperatur. Die Nullhypothese „normale Körpertemperatur" scheint also zutreffend. Die Körpertemperatur von B hingegen liegt bei 37,7 °C und damit außerhalb der Streubreite der normalen Temperatur. Nur 2,5 % der gesunden Menschen haben von Natur aus eine derart hohe Körpertemperatur. Die Nullhypothese „normale Körpertemperatur" scheint hier also nicht zutreffend. Das Gleiche lässt sich für die Körpertemperatur von C sagen, denn nur 2,5 % der gesunden Menschen haben von Natur aus eine Körpertemperatur unter 36,3 °C. Die Nullhypothese „normale Körpertemperatur" scheint also auch hier nicht zutreffend. Sie schließen daraus, dass B und C wahrscheinlich krank sind.

Doch genau hier liegt das Problem: Sie *könnten* falsch liegen. Wenn B und C tatsächlich rein zufällig zu den 2,5 % der Menschen gehören, die von Natur aus eine so hohe bzw. niedrige Körpertemperatur haben, würde dies bedeuten, dass Sie eine wahre Nullhypothese zurückweisen. Eine solch falsche Schlussfolgerung bezeichnet man auch als *Fehler 1. Art* oder *falsch-positive Entscheidung*. Die Wahrscheinlichkeit, dass Sie diese Art von Fehler machen, liegt bei 2,5 %, denn 2,5 % der gesunden Menschen haben eine Körpertemperatur, die tatsächlich so hoch (niedrig) oder höher (niedriger) liegt. Demnach liegt die Wahrscheinlichkeit, dass Sie einen falschen Schluss gezogen haben, bei 2,5 von 100. Man spricht auch von einem Alpha-Fehler (α-Fehler, Annahmefehler) und bezeichnet damit die Wahr-

Tab. 7.1 Die Logik der Hypothesentests.

Entscheidung	**Wahrheitsgehalt**	
	H_0 wahr	**H_0 falsch**
Ablehnung von H_0 Entscheidung: Der Patient hat eine normale Temperatur.	Fehler 1. Art	richtige Entscheidung
Nicht-Ablehnung von H_0 Entscheidung: Der Patient hat eine vom Normalwert abweichende Temperatur.	richtige Entscheidung	Fehler 2. Art

scheinlichkeit, die richtige Annahme der Nullhypothese (H_0) falscherweise zurückzuweisen. Ein Alpha-Fehler kann auch als quantitatives Maß für die Ergebnisunsicherheit bei häufiger Wiederholung der Stichproben betrachtet werden.

Es gibt einen weiteren Fehler, den Sie bei der Überprüfung einer Hypothese machen können, nämlich den, eine falsche Nullhypothese nicht zurückzuweisen. Für unser Beispiel bedeutet dies, dass Sie entscheiden, 37,7 °C ist für Person B die normal hohe Temperatur und daraus folgern, Person B hat kein Fieber. Doch nehmen wir einmal an, Sie liegen falsch, und Person B ist tatsächlich krank. Dies bedeutet, dass Sie eine falsche Nullhypothese nicht zurückzuweisen. Dieser Fehler wird als *Fehler 2. Art* oder *falsch-negative Entscheidung* bezeichnet. Man spricht auch von einem Beta-Fehler (β-Fehler) und bezeichnet damit die Wahrscheinlichkeit, die falsche Annahme der Nullhypothese (H_0) beizubehalten und nicht zurückzuweisen. Tabelle 7.1 stellt dies grafisch sehr klar und anschaulich dar.

Der springende Punkt ist, dass man im Rahmen wissenschaftlicher Untersuchungen die Wahrheit nie mit absolu-

ter Sicherheit wissen kann. Man ist grundsätzlich immer auf Schätzungen und Annahmen angewiesen, die auf verfügbaren Informationen basieren. Um dies zu kompensieren, empfiehlt es sich, die Sicherheit über den Wahrheitsgehalt einer Hypothese zu maximieren, indem man die Wahrscheinlichkeit fehlerhafter Schlüsse bestmöglich minimiert. Eine wahre Nullhypothese zurückzuweisen, ist für gewöhnlich ein schwerer Fehler, weshalb man gemeinhin darauf zielt, die Wahrscheinlichkeit für den Alpha-Fehler bei 5 % und damit sehr gering zu halten. Wenn Ihre Nullhypothese als einseitiger Test konzipiert ist, wenden Sie die vollen 5 % auf den obersten Wert der Verteilung an. Wenn Ihre Nullhypothese als zweiseitiger Test konzipiert ist, sind Sie an zwei Punkten der Verteilungswerte interessiert – und zwar einmal an dem, der außerhalb der markierten Obergrenze von 2,5 % liegt, und einmal an dem, der außerhalb der markierten Untergrenze von 2,5 % liegt. Diese Bereiche werden als „kritische Bereiche" für die Zurückweisung einer Nullhypothese bezeichnet (da sie eine Anzahl „kritischer Werte" anzeigen, von denen wir nicht mehr glauben können, es seien reine Zufallstreffer).

All dies im Hinterkopf entscheiden Sie sich nun, die gleiche Logik auf die Depressionswerte Ihrer beiden Gruppen anzuwenden (unter der Annahme, dass sich diese Depressionswerte ebenfalls in einer Glockenkurve verteilen). Der mittlere Depressionswert der Placebogruppe kann als „normaler" oder „erwarteter" Wert angenommen werden, ebenso wie 37 °C als „normale" oder „erwartete" Körpertemperatur gilt. Sie vergleichen nun den mittleren Depressionswert in der behandelten Gruppe mit dem mittleren Depressionswert in der Placebogruppe. Zur Berechnung

der Wahrscheinlichkeitsverteilung, dass die beobachteten unterschiedlichen Werte zwischen beiden Gruppen innerhalb der normalen Streubreite (Variabilität) liegen, brauchen Sie als Maßzahl die jeweilige Gruppengröße sowie die Varianz, also das Maß für die mittlere Abweichung der Werte innerhalb jeder Gruppe. Sie können die Nullhypothese nur dann zurückweisen („kein Unterschied"), wenn die Menge aller Ergebnisse innerhalb des Ablehnungsbereichs (und damit im kritischen Bereich) liegt. Wenn Sie im oberen Ablehnungsbereich liegen, können Sie schließen, dass Ihr SSRI-Medikament die Depression erhöht. Wenn Sie im unteren Ablehnungsbereich liegen, können Sie schließen, dass Ihr SSRI-Medikament die Depression lindert.

Nehmen wir an, Sie vergleichen beide Gruppen miteinander und kommen zu dem Schluss, dass sich die Ergebnisse ziemlich unterschiedlich darstellen. Tatsächlich ist der Unterschied so groß, dass die Ergebniswerte allesamt in einen der Ablehnungsbereiche fallen. Nehmen wir weiterhin an, dass die Wahrheit darin liegt, dass Ihr SSRI-Medikament *nicht* wirkt. Der Unterschied ist auf Zufallsfaktoren oder Messfehler zurückzuführen. Sie lehnen die Nullhypothese ab und begehen damit einen *Fehler 1. Art*, denn Sie haben die Nullhypothese zurückgewiesen, obwohl sie wahr ist. Ihr SSRI-Medikament wirkt nicht, aber Sie schließen dennoch, *dass* es wirkt.

Nehmen wir nun stattdessen an, Sie vergleichen beide Gruppen miteinander und kommen zu dem Schluss, dass sich die Ergebnisse ziemlich gleich darstellen. Einen kleinen Unterschied gibt es zwar, doch der ist so gering, dass die Werte nicht in die kritischen Ablehnungsbereiche fallen. Sie entscheiden daher, dass Ihr SSRI-Medikament

nicht wirkt. Aber nun nehmen wir an, dass die Wahrheit darin liegt, *dass* Ihr SSRI-Medikament *wirkt*. Es erzeugt einen geringen, aber wahren Unterschied zwischen beiden Gruppen, und vielleicht bewirken höhere Dosen oder eine Behandlung über einen längeren Zeitraum einen wirklich signifikanten Unterschied. Sie weisen die falsche Nullhypothese nicht zurück und begehen damit einen *Fehler 2. Art.*

Nehmen wir an, dass Sie genau dies tatsächlich beobachten. Ihre Ergebnisse fallen nicht in die Ablehnungsbereiche. Statistisch gesehen können Sie die Nullhypothese also nicht zurückweisen. Aber schließen Sie daraus, dass Ihr SSRI-Medikament nicht wirkt? Ist dies der Schluss, den ein Wissenschaftler tatsächlich ziehen würde?

Wie wir gleich sehen, brauchen wir gar nicht zu spekulieren. Dieses Beispiel basiert auf einer konkreten Fallstudie. Ein Forscherteam um Irvin Kirsch von der University of Hull (2008) konnte sämtliche klinische Studien einholen, die der behördlichen Lebensmittelüberwachungs- und Arzneimittelzulassungsbehörde der USA (Food and Drug Administration, FDA) zur Lizenzierung von vier Antidepressiva (selektiven Serotoninwiederaufnahmehemmern, SSRI) vorlagen – Prozac, Effexor, Paxil und Serzone. Die Forscher fanden heraus, dass diese Arzneien bei Patienten mit schwerer Depression sehr wohl wirken, nur eben nicht besser als Scheinmedikamente (Placebos) bei leichter oder mittelschweren Depression. Sie schlossen daraus: „Es geht den Patienten nach der Einnahme von Antidepressiva durchaus besser, aber das ist auch nach der Einnahme von Placebos der Fall. Ein signifikanter Unterschied hinsichtlich der Verbesserung ist also nicht feststellbar. Dies bedeutet, dass depressive Menschen eine Besserung auch

ohne biochemische Behandlung erreichen können." Mary Ann Rhyne, Sprecherin der Herstellerfirma von Paxil, widersprach diesem Schluss, der ausschließlich auf klinischen Studien der FDA basierte. „Die Autoren", so kritisierte sie, „negieren den überaus positiven Nutzen, den die Behandlung mit Antidepressiva betroffenen Patienten und ihren Familien bringt, und stehen damit in einem auffälligen Widerspruch zu den Erfahrungen der aktuellen klinischen Praxis." Beachten Sie bitte, dass Rhynes Kritik auf einem echten Gegensatz beruht, auf dem, was Mediziner „sehen", und dem, was wissenschaftliche Daten zeigen.

Box 7.2 Statistik oder Bayes?

Wie wir in Kap. 3 gesehen haben, gibt es häufig eine strikte Trennung zwischen wissenschaftlichen Studien und klinischen Entscheidungen. Ein Arzt möchte wissen, wie hoch die Wahrscheinlichkeit ist, dass sein Patient eine bestimmte Erkrankung hat. Ein Wissenschaftler will wissen, ob ein bestimmter Faktor die Ursache einer Erkrankung ist. Der Arzt muss eine „bayesische" Entscheidung treffen; der Wissenschaftler muss ein Experiment durchführen und eine Annahme auf ihre statistische Signifikanz überprüfen, um Nullhypothesen zurückweisen zu können.

Das folgende Beispiel führt uns sehr schön vor Augen, wie diese unterschiedlichen Ansätze auch in der Wissenschaft zu unterschiedlichen Antworten führen (aus Anderson 1998):

Mal angenommen, ein Ökologe sucht in einem Wald nach Alkenvögeln. Welche Anzeichen deuten wohl darauf hin, dass Alkenvögel dort zu finden sind? Voraussetzungen für geeignete Nistplätzen sind ein Kriterium: Bäume mit großen, ausladenden Ästen, dicht überwachsen mit Moos und Flechten. Der Forscher macht eine standardisierte Be-

standsaufnahme des Waldes, um solche Äste auszumachen, hat 1000 solcher Bäume ausgemacht und folgende Daten ermittelt:

	H_0 wahr: Alkenvögel nisten *nicht* in diesem Baum	**H_0 falsch: Alkenvögel nisten in diesem Baum**	**insgesamt**
Untersuchungsdaten: *keine* potenziellen Nistplätze	808	4	812
Untersuchungsdaten: potenzielle Nistplätze gefunden	142	46	188
insgesamt	950	50	1000

H_1: geeignete Nistplätze werden mit dem Vorkommen nistender Alkenvögel assoziiert

H_2: geeignete Nistplätze werden *nicht* mit dem Vorkommen nistender Alkenvögel assoziiert

Die nach Hypothese H_0 zu erwartenden Häufigkeiten der einzelnen möglichen Beobachtungen liegen bei:

$$(50 \times 812)/1000 = 41$$

$$(50 \times 188)/1000 = 9$$

$$(950 \times 812)/1000 = 771$$

$$(950 \times 812)/1000 = 179$$

Wir setzen diese zusammen mit den tatsächlich auftretenden Häufigkeiten der jeweiligen Beobachtungen in die Formel für die Prüfgröße χ^2 ein und erhalten:

$$\chi^2 = (4 - 41)^2 / 41 + (46 - 9)^2 / 9 + (808 - 771)^2 / 771 + (141 - 179)^2 / 179 = 186$$

Die Wahrscheinlichkeit, einen *Fehler 1. Art* zu begehen, wenn H_0 zurückgewiesen wird, liegt bei weniger als 0,0001! Geeignete Nistplätze sind also sicherlich gute Orte, um Alkenvögel zu finden.

Aber wie hoch ist die Wahrscheinlichkeit, dass Alkenvögel auch tatsächlich an diesen Orten nisten, nur weil unser Forscher die Bäume als potenzielle Nistplätze ausgemacht hat? Ganz einfach: 46/188=0,24. Sind diese Orte ein guter Indikator für Nistplätze? Nein. Die Chance, in diesen Bäumen nistende Alkenvögel zu finden, liegt bei weniger als 1 von 4 (25 %). In einem medizinischen Kontext bedeutet dies, dass für manche Fragestellungen das „bayesische" Denken angemessener ist („Wie wahrscheinlich ist es, dass dieser Patient diese Krankheit hat?"), für andere wiederum sind statistische Signifikanztests geeigneter („Ist dieser Faktor eine wahrscheinliche Ursache für diese Krankheit?").

Backup-Systeme für alle Fälle

Wie dieses Kapitel zeigt, sind Hypothesentests hochkomplexe Verfahren voller Verzerrungen, die irrational sein können oder nicht. Aber heißt dies, dass wir sie als wissenschaftliche Untersuchungsmethode aufgeben sollten?

Um die Wahrscheinlichkeit für Fehler (Fehlentscheidungen) bei Hypothesentests möglichst gering zu halten, verfügt die Wissenschaftsgemeinde über etliche Backupsys-

teme. Zu einen ist es unter Wissenschaftlern gängige Praxis, Experimente zu wiederholen und zu überprüfen, um Ergebnisse zu duplizieren. Das Experiment, das uns hier als Beispiel diente, haben unabhängige Forschungslabors in sogenannten Replikationsstudien unter gleichen wie auch unter veränderten Versuchsbedingungen wiederholt. Eine Replikation ist ein überaus wichtiges Verfahren, um wissenschaftliche Irrtümer aufzuspüren. Zum anderen unterliegen die Methoden und Verfahren, die in wissenschaftlichen Studien durchgeführt werden, gegenseitigen Begutachtungs- und Kontrollprozessen (*peer review*). Wissenschaftler senden Manuskripte über ihre Arbeit an Fachmagazine. Die Herausgeber der Magazine senden die eingereichten Manuskripte weiter an Fachgutachter, die mit dem jeweiligen Gebiet vertraut sind. Die Gutachter können ein Manuskript unverändert oder mit Änderungsvorschlägen zur Veröffentlichung empfehlen oder es ablehnen, wenn es von nicht ausreichender Qualität oder Bedeutung ist. Die Ablehnungsrate durch die Wissenschaftsverlage ist sehr hoch und liegt mitunter bei 95 %.

Aber liegen die Gutachter immer richtig? Wahrscheinlich nicht. Wissenschaftler sind auch nur Menschen. Gewiss, Gutachten verhindern in vielen Fällen, dass qualitativ schlechte Arbeiten veröffentlicht werden, es kommt aber auch vor, dass Vetternwirtschaft oder potenzielle Gewinnaussichten eine Veröffentlichung befördern. Hin und wieder aber führen fundierte Studien, deren Ergebnisse nicht mit geltenden Theorien erklärt oder mit anerkannten Methoden untersucht werden können, auch zu echten Umbrüchen im wissenschaftlichen Diskurs über ein bestimmtes Phänomen. Der Philosoph Thomas Kuhn (1922–1996)

vertritt in seinem 1962 erschienen Buch *Die Struktur wissenschaftlicher Revolutionen* die These, dass wissenschaftlicher Fortschritt sich nicht durch linear verlaufende Prozesse der Anhäufung neuer Erkenntnisse vollzieht, sondern in unregelmäßig periodischen Abständen gravierende Umwälzungen durchläuft, in denen herkömmliche Annahmen und Denkweisen über ein Phänomen durch vollkommen neue abgelöst werden. Durch diese Umwälzungen kommt es zu „revolutionären Paradigmenwechseln", wie Kuhn es bezeichnet: Alle alten Hypothesen einer Disziplin werden über den Haufen geworfen und durch neue ersetzt.

In seinen Studien über die Vorgehensweise der Wissenschaft, stellte Kuhn fest, dass wissenschaftlicher Fortschritt aus drei unterschiedlichen Phasen besteht. Die erste bezeichnet er als Vor-Wissenschaft; ein bestimmtes Phänomen erregt die Aufmerksamkeit eines Wissenschaftlers oder einer Gruppe von Wissenschaftlern, und sie beginnen, sich damit zu beschäftigen, um es zu erklären und zu verstehen. In dieser Phase fehlt ein zentrales Paradigma insofern, als dass noch niemand weiß, wie das Phänomen begrifflich zu fassen ist. Es folgt die zweite Phase, die „ normale Wissenschaft". Nachdem sie genügend Experimente durchgeführt haben, meinen die Wissenschaftler, das Phänomen zu verstehen und postulieren eine Theorie, um es zu erklären. Diese Theorie wird zum zentralen Paradigma, zu einer Festlegung darüber, wie das Phänomen zu begreifen ist. Doch irgendwann wird ein nicht bestätigendes Ergebnis berichtet. Dieses Ergebnis jedoch wird typischerweise nicht als eine Widerlegung des bestehende Paradigmas erachtet, sondern als ein Fehler oder Irrtum des Wissenschaftlers. Doch mit einer zunehmenden Häufung von widerlegenden

Ergebnissen gerät die Wissenschaft in eine „Krise“ und tritt damit in die dritte Phase ein. An diesem Punkt schlägt normalerweise irgendwer ein völlig anderes theoretisches Rahmenkonzept vor, um das Phänomen in einer anderen Weise zu erklären und die bis dahin problematischen Ergebnisse darin einzupassen. Das neue Paradigma löst das alte ab, je mehr Wissenschaftler es als gültig annehmen. Kuhn nennt dies „revolutionäre Wissenschaft“.

Dabei ist wichtig, dass während der Umbruchphase konkurrierende Paradigmen existieren, die miteinander unvereinbar sind; sie erklären das Phänomen in unterschiedlicher Weise und ziehen häufig Konzepte heran, die nur zu einer theoretischen Sichtweise passen. Kuhn stellte des Weiteren fest, dass Paradigmen abgelöst werden, wenn fünf Faktoren erfüllt sind: Das neue Paradigma muss

- in seinen Prognosen genauer als das alte,
- in sich kohärent und mit andern Theorien konsistent sein,
- umfassender angelegt sein als das alte, da es eine größere Vielzahl von Effekten erklärt,
- einfacher sein als das alte, und
- dem Fortschritt dienlich sein, da es neue Prognosen trifft, die das alte nicht getroffen hat.

Die Geschichte der Wissenschaft ist voll von dieser Art von Paradigmenwechseln. Seit Jahrtausenden galt das Paradigma der Himmelsmechanik von Aristotes und Ptolemäus, wonach sich die Sonne und andere Gestirne um die Erde drehen. Doch dann veröffentlichte Nikolaus Kopernikus 1543 seine Theorie vom Umlauf der Planeten um

die Sonne. Damit vertrat er eine Sicht, die nicht nur das von Aristoteles bzw. Ptolemäus geprägte Weltbild bedrohte, sondern auch der klerikalen Lehre widersprach, wonach die Erde der Mittelpunkt des Universums war. Der Vatikan schritt ein und setzte Kopernikus' Buch auf die Liste der verbotenen Bücher. Danach blieb es 80 Jahre lang unbeachtet, bis der italienische Wissenschaftler Galileo Galilei (1564–1642) sein eigenes, neu erfundenes Teleskop erstmals gen Himmel richtete und überzeugende Beweise für das kopernikanische Weltbild fand. Er erkannte, dass die Venus genau wie der Mond bestimmte Phasen durchläuft (voll, halb, abnehmend, zunehmend). Dies war nur möglich, wenn sich die Venus um die Sonne bewegte, nicht um die Erde. Er erkannte auch, dass die Monde des Jupiters um den Jupiter kreisen, nicht um die Erde. Das rief den Vatikan erneut auf den Plan, und man verlangte von Galilei, seine neuen Lehren zu widerrufen, andernfalls drohe ihm die Exkommunikation oder gar die Hinrichtung. Galilei kam der Aufforderung nach, doch tief in seinem Inneren wehrte er sich, sich dem Zwang zu beugen und diese so bedeutsame Wahrheit zu verleugnen. Er verfasste und veröffentlichte seinen berühmten *Dialog* (1632), in dem er die kopernikanische Lehre verteidigte und unterstützte und durch einen Vergleich das alte ptolemäische Weltbild als unhaltbar bewies. Der Vatikan reagierte prompt, forderte ihn unter Androhung von Folter auf, von seinen absurden Theorien abzuschwören. Als er sich weigerte, wurde er zu einer lebenslangen Haftstrafe verurteilt, die später gemildert und in einen lebenslangen Hausarrest umgewandelt wurde. Doch manche Ideen gehen einfach nicht unter. Andere Wissenschaftler (vor allem der deutsche Mathematiker

und Philosoph Johannes Kepler, der dänische Astronom Tycho Brahe und der englische Naturforscher Isaac Newton) erkundeten den Himmel weiterhin und verteidigten letztlich die kopernikanische Sicht.

Zu den wissenschaftlichen Revolutionen der neueren Zeit gehören Einsteins Relativitätstheorie oder frühe Quantentheorie, die die klassische newtonsche Physik ablösten, oder auch die Bakteriologie, die die Miasmentheorie widerlegte, sowie die Keimplasmatheorie, die das Paradigma der Pangenesistheorie ablöste. Auch die Psychologie verzeichnet in ihrer relativ jungen Geschichte bereits etliche Paradigmenwechsel, da zunehmende Fortschritte in den Computerwissenschaften immer mehr Erkenntnisse über kognitive Vorgänge im Gehirn möglich machten. Im ausgehenden 19. Jahrhundert sahen frühe Vertreter der strukturalistischen Psychologie ihre Rolle darin, mentale Ereignisse wie Gefühle und Gedanken durch Introspektion oder Einfühlung zu untersuchen und zu erklären. Dieser Ansatz wurde im frühen 20. Jahrhundert durch Erkenntnisse der behavioristischen Forschung gewaltig erschüttert und umgestoßen. Die Vertreter des Behaviorismus (der Wissenschaft vom Verhalten) waren der festen Überzeugung, dass echte Wissenschaft auf direkt beobachtbaren Vorgängen wie beispielsweise dem Verhalten gründen müsse. Ungefähr zur gleichen Zeit begannen Computerwissenschaftler Maschinen zu erschaffen, die intelligente Funktionen ausführen konnten, und begründeten damit einen neuen Zweig, der in den 1960er-Jahren seine Blüte erreichte. Was diese intelligenten Programme tun und wie sie arbeiten, ließ sich nur beschreiben unter Bezug auf systeminterne Zustände, Datenstrukturen und Zielgrößen. Daraus leiteten die Psy-

chologen eine ganz neue Fragestellung ab – „Wenn eine Maschine innere Zustände haben kann, warum nicht auch der Mensch?“ – und nahmen damit den Behavioristen das Heft wieder aus der Hand. Die kognitive Psychologie war geboren und floriert bis heute.

Wie können wir diesen wissenschaftlichen Prozess zusammenfassen? Trotz der Paradigmenwechsel und wissenschaftlichen Revolutionen ist und bleibt das werktägliche Kernelement einer jeden wissenschaftlichen Untersuchung die hypothetisch-deduktive Methode. Wissenschaftler postulieren Hypothesen, die sie aus Theorien ableiten, und prüfen sie nach Poppers Modell der scFalsifikation. Diese Methode zum Nachweis der Ungültigkeit einer Aussage, ist in der Wissenschaftsgemeinde weithin akzeptiert, da es die beste ist, die es derzeit gibt. Sie ist folgendermaßen formuliert:

- Die beste wissenschaftliche Theorie ist die, die empirisch geprüft werden kann.
- Experimente müssen wiederholbar und nachprüfbar sein und den größten zu erwartenden Informationsgewinn erbringen.
- Die Erklärungen und Theorien der wissenschaftlichen Disziplinen basieren auf einem breiten Konsens darüber, dass sie durch Beobachtungen wiederholt bestätigt werden können, widerspruchsfrei sind und mit der empirischen Wirklichkeit übereinstimmen.

Da sich gegenwärtige wissenschaftliche Theorien gerne auf den vielzitierten *breiten Konsens* stützen, schlossen einige Gelehrte, dass wissenschaftliche Theorien nichts weiter

seien als kulturelle Konventionen und keine wahren Beschreibungen der Realität (für eine Kritik postmoderner Analysen, s. Spiro 1996). Das wiederum stößt vielen anderen sauer auf, denn das wäre so, als würde man schließen, dass es so etwas wie Elefanten nicht gibt, nur weil die sprichwörtlichen Blinden, die ihn untersuchen, allesamt zu unterschiedlichen Beschreibungen gelangen („Ein Elefant ist so was wie ein Strick", sagt der Blinde, der seinen Rüssel fühlt; „Ein Elefant ist so was wie ein Baum", sagt der Blinde, der seine Beine fühlt usw.) Hypothesentests gelten heute als eine der beweiskräftigsten Methoden, um die Wahrheit einer Aussage zu belegen oder zu widerlegen. Daniel Bernoulli (1700–1782) beispielsweise verwendete wissenschaftliche Hypothesentests, um zum ersten Mal überhaupt das Prinzip des Auftriebs und damit die Physik des Fliegens zu erklären. Und wie hat der altkluge, sechsjährige Bertie, der kleine Held in *Love over Scotland* von Alexander McCall Smith (2007), einmal gesagt? Ohne Vertrauen in die Zuverlässigkeit und Richtigkeit wissenschaftlicher Schlussfolgerungen würde wohl kaum einer ein Flugzeug besteigen.

8

Problemlösungen

Vom problemorientierten zum lösungsorientierten Denken

„Houston, wir haben ein Problem." Wir schreiben das Jahr 1970 und die NASA hat mit Apollo 13 soeben eine weitere bemannte Expedition in den Weltraum geschickt. Es ist die siebte bemannte Expedition und die dritte, für die eine Mondlandung geplant ist. Die Menschen sind mittlerweile ziemlich gewöhnt an diese Expeditionen, sodass kaum einer mehr die Mission vor den Fernseher verfolgt … bis ein Sauerstofftank explodiert, das Raumschiff beschädigt und die drei Astronauten in 330.000 km über der Erde einem Kampf um Leben und Tod aussetzt. Während die Astronauten in der begrenzten Zeit, die ihnen bleibt, nach Kräften um ihr Überleben kämpfen, sucht die heroische Crew im Kontrollzentrum der NASA in Houston nach Mitteln und Wegen, die Besatzung mit den an Bord verfügbaren Systemen sicher nach Hause zu bringen. Wie wir alle aus der Geschichte wissen (und aus einer großartigen Verfilmung unter der Regie von Ron Howard mit Tom Hanks in der Hauptrolle), finden sie eine Lösung und das Drama findet ein glückliches Ende.

Nun ist die Geschichte um Apollo 13 in ihrer Dramatik nicht zu vergleichen mit den Problemen, denen wir immer mal wieder in unserem Alltag begegnen, doch sie enthält

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_8,

beispielhaft alle Zutaten, die nötig sind, um Probleme zu überwinden und sie letztlich zu lösen – seien es kleine lästige Probleme des alltäglichen Lebens (verlegte Schlüssel) oder potenziell lebensverändernde Technologien (ein neues Heilmittel gegen eine tödliche Krankheit). Die Lektion, die wir aus der Katastrophe von Apollo 13 (wie auch aus den Ergebnissen aus 100 Jahren problemorientierter Forschung) ziehen, ist die: *Problemlösen heißt Suche nach Erkenntnis und Information.* Wenn Sie Ihre ganze Bude auf den Kopf stellen, weil Sie Ihre Schlüssel verloren haben, haben Sie sich auf die Suche gemacht. Wenn Ermittlungsbeamte einen Tatort durchkämmen, gehen sie auf die Suche nach Beweisen, um das Verbrechen aufzuklären. Wenn ein Staatsanwalt einen Zeugen befragt, sucht er nach erinnerten Informationen, die zur Aufklärung des Verbrechens führen. Wenn ein Mediziner Laborexperimente mit Krebszellen durchführt, sucht er nach Mechanismen, die einer Krebserkrankung zugrunde liegen. Wenn Sie über ein Problem grübeln, durchforsten sie ihren Wissensspeicher nach Fakten und Zusammenhängen, die zu einer Lösung führen. Und wenn Sie mitten in der Nacht aus dem Schlaf fahren, weil Sie die Lösung für ein Problem gefunden haben, das Sie den ganzen Tag lang beschäftigt hatte, so liegt dies daran, dass Ihr Gehirn weiterhin aktiv war, innere Suchprozesse in Gang gesetzt und die Lösung schließlich gefunden hat.

Problemlösung bedeutet also die Suche nach Informationen. Diese scheinbar simple Erkenntnis hat allerdings weitreichende Folgen, nicht nur was die Verbesserung von Lösungsstrategien angeht, sondern auch im Hinblick auf automatisierte Abläufe. Wir verlassen uns heutzutage regelmäßig auf automatisierte Systeme zur Lösung unserer Pro-

bleme. Automatisierte Systeme rufen relevante Informationen aus dem Internet für uns ab, bewegen Roboterarme, die als Assistenten bei medizinischen Operationen oder im Autobau fungieren, und sie betreiben sogar Aktienhandel. Und das alles, weil wir eine Menge darüber gelernt haben, wie wir Suchprozesse beschreiben und implementieren. Und wir haben eine Menge darüber gelernt, wie wir Probleme kennzeichnen und charakterisieren.

Karl Duncker (1903–1940), Psychologe und Mitbegründer der Gestalttheorie, schreibt darüber in seinem berühmten Aufsatz (1945):

> „Ein ‚Problem' entsteht z. B. dann, wenn ein Lebewesen ein Ziel hat und nicht‚weiß', wie es dieses Ziel erreichen soll. Wo immer sich der gegebene Zustand nicht durch bloßes Handeln (Ausführen selbstverständlicher Operationen) in den erstrebten Zustand überführen lässt, wird das Denken auf den Plan gerufen. Ihm liegt es ob, ein vermittelndes Handeln allererst zu konzipieren."[1]

Insbesondere im Jahr 1945 wurden viele neue, wissenschaftliche Erkenntnisse auf dem Gebiet der Problemlösung publiziert. Aufsehen erregten vor allem die grundlegenden Schriften des ungarischen Mathematikers George Pólya, die unter dem passenden Titel *How to solve it* erschienen (deutsch: „Schule des Denkens. Vom Lösen mathematischer Probleme"). Pólya fasste Problemlöseprozesse in vier Phasen zusammen. Wie genial ihm das gelang, sei hier kurz dargestellt:

[1] Duncker K (1935) Zur Psychologie des produktiven Denkens. Springer, Berlin

1. Verstehen der Aufgabe: Machen Sie sich die wesentlichen Aspekte des Problems klar. Durchsuchen Sie Ihren Wissensspeicher nach verwandten Problemen.
2. Ausdenken eines Plans: Durchsuchen Sie Ihren Wissensspeicher nach Informationen, die Ihnen lösungsrelevant erscheinen.
3. Ausführen des Plans: Setzen Sie Ihre Lösungen um.
4. Rückschau: Reflektieren Sie unter dem Aspekt „Was hätte ich besser machen können?"

Wie Duncker klar formuliert, entsteht ein Problem dann, wenn ein bestehender Zustand (Ausgangs- oder Istzustand) sich nicht mit einem gewünschten Zielzustand deckt. Können wir ein Problem nicht auf Anhieb lösen, um das gewünschte Ziel zu erreichen, setzen normalerweise (produktive) Denkprozesse ein. Ziel dieser Denkprozesse ist es, wie sowohl Duncker als auch Pólya betonen, Mittel und Wege (oder Handlungen) zu finden, die wir ergreifen können, um dem gewünschten Ziel näher zu kommen. Der Prozess des Problemlösens lässt sich daher als eine sogenannte *Differenzreduktion* beschreiben – als die Suche nach einem Operator, der die Differenz zwischen dem gegebenen und dem Zielzustand (maximal) reduziert.

Stellen Sie sich vor, Sie wollen ins Kino gehen und können Ihre Autoschlüssel nicht finden. Ihr Zielzustand, das Ziel, ist „Kino". Ihr aktueller Zustand ist „zu Hause". Ihr Problem: Sie befinden sich in einer Situation, in der Ihr Zielzustand vom Ausgangszustand abweicht. Sie versuchen daher alles Mögliche, um ins Kino zu kommen. Vielleicht rufen Sie einen Freund an, damit er Sie abholt. Oder Sie gehen zu Fuß, sofern das Wetter schön und das Kino

nicht allzu weit weg ist. In der Fachsprache der Problemlöser spricht man hier von *Mittel* oder *Operatoren*, die Sie einsetzen können, um die Differenz zwischen Ziel- und Ausgangszustand zu verringern. Und Sie wissen, dass Ihr Problem erst dann behoben ist, wenn Ihr Ausgangszustand (im Kino sitzen) Ihrem gewünschten Zielzustand (im Kino sitzen) entspricht. Problem gelöst!

Und was, wenn sich unmittelbar keine praktische Lösung für ein Problem finden lässt? Genau dann bemühen wir unsere Denk- und kreative Vorstellungskraft.

Wenn Probleme klar definiert sind

Manche Probleme sind klar definiert – das heißt, sie haben jeweils genau festgelegte Ausgangs-(Ist-) und Ziel-(Soll-)zustände wie auch Lösungswege. Es ist ganz einfach, den Weg ins Kino zu finden, wenn man nur in die richtige Richtung denkt; es ist auch ganz einfach, die Lösung von zwölf geteilt zu vier auszurechnen, wenn man die mathematischen Divisionsregeln kennt; und es ist ebenso einfach, Rühreier zuzubereiten, wenn man ein gutes Rezept hat.

Das Schöne an klar definierten Problemen ist, dass es meist Algorithmen gibt, um sie zu lösen. Ein Algorithmus ist ein Vorgang, der nach einem bestimmten Schema abläuft und ein Ergebnis hat – wie etwa, systematisch im ganzen Haus, Zimmer für Zimmer, nach den verlorenen Schlüsseln zu suchen. Früher oder später kommen Sie ins letzte Zimmer (und der Vorgang wird ein Ergebnis haben), und sollten sich die Schlüssel im Haus befinden, werden Sie sie auch garantiert finden. Wenn nicht, ist der Vorgang

ebenfalls beendet, denn Sie sind ja im letzten Zimmer angekommen. Die meisten Algorithmen für klar definierte Probleme garantieren nicht nur den Abschluss mit einem Ergebnis, sie garantieren auch das richtige Ergebnis. Wenn Sie die Regeln der Multiplikation genau befolgen, werden Sie die korrekte Antwort für das Problem „296 × 398" in einer endlichen Zeit finden. Und wenn Sie sich genau an ein Rezept für „Rührei" halten, wird das Rührei in einer endlichen Zeit auch fertig sein.

Lassen Sie uns diese Problemlösungsmethode etwas präziser beleuchten: *Ein Algorithmus besteht aus einer Reihe endlich vieler, wohldefinierter Einzelschritte, die ein bestimmtes Problem lösen*. Jeden einzelnen Begriff in diesem Satz können wir definieren, und zwar so:

- Einzelschritt: Jede einzelne Handlung, die Sie auf dem Weg zu einem Ergebnis ausführen müssen, wird als Schritt bezeichnet. Im Fall des Rühreirezepts besteht ein Schritt im Aufschlagen der Eier.
- Reihe: Alle Schritte müssen in einer bestimmten Reihenfolge erfolgen und jeder dieser Schritte muss ausgeführt werden (es sei denn, der Algorithmus gibt etwas anderes vor). Sie müssen also zuerst die Eier aufschlagen, bevor Sie sie verrühren.
- wohldefiniert: Ein Schritt *darf nicht* durch einen ähnlichen Schritt ersetzt werden. Das heißt: Sie müssen die Eier aufschlagen und dürfen sie nicht zerschlagen.
- lösen: Ein Algorithmus erzeugt ein abschließendes Ergebnis (Endergebnis). Wenn Sie sich genau an das Rezept halten, haben Sie am Ende fertig gebackene Rühreier.

- bestimmtes Problem: Der Algorithmus für ein Problem löst zumeist nur dieses eine Problem, nicht auch ein anderes. Ihr Rezept gilt speziell für Rühreier. Wenn Sie einen Kuchen backen wollen, müssen Sie einen anderen Algorithmus (sprich, ein anderes Rezept) finden.

Eine besonders effiziente Form der Algorithmen sind *rekursive* Algorithmen. Vorgänge sind *rekursiv,* wenn sie zwei Bedingungen erfüllen: Es muss einen Basisfall geben und die Regeln müssen alle anderen Fälle systematisch an den Basisfall annähern. Einfacher ausgedrückt: *Eine rekursive Vorgehensweise wendet Regeln an, die sich auf sich selbst beziehen*. Damit dies funktioniert, müssen auch die folgenden Bedingungen gelten: Das Problem muss lösbar sein und es muss einen abschließenden Satz geben (das heißt, man muss wissen, wann man fertig ist).

Ein klassisches Beispiel für ein Problem, das sich mit einem rekursiven Algorithmus lösen lässt, sind die sogenannten „Türme von Hanoi", ein mathematisches Knobelspiel (Box 8.1). Zu Beginn des Spiels stapeln Sie mehrere kreisrunde, gelochte Scheiben mit unterschiedlichem Durchmesser auf einen von drei Stäben, und zwar auf den Stab, der sich links außen befindet (Stab A). Die größte Scheibe legen Sie zuunterst und türmen darauf dann der Größe nach geordnet alle anderen, bis die kleinste Scheibe zuoberst liegt. Das Ziel ist nun, alle Scheiben von Stab A links außen auf Stab C rechts außen zu versetzen, sodass am Ende wieder die größte Scheiben unten und die kleinste oben liegt. Der Kniff dabei ist, dass immer nur eine Scheibe bewegt werden darf und immer und überall eine kleinere auf einer größeren Scheiben liegen muss. Klingt einfach, aber probieren Sie es einmal selbst! Es ist zum Verrücktwerden!

Box 8.1 Die „Türme von Hanoi"

Die „Türme von Hanoi" sind ein klassisches Beispiel für die Art von Problemen, die sich mithilfe einer rekursiven Vorgehensweise lösen lassen. Wie Sie aus der Abbildung ersehen können, gibt es drei Stäbe. Ihre Aufgabe besteht darin, die Scheiben einzeln und nacheinander von A nach C zu bewegen, ohne dass eine größere Scheibe auf einer kleineren zu liegen kommt.

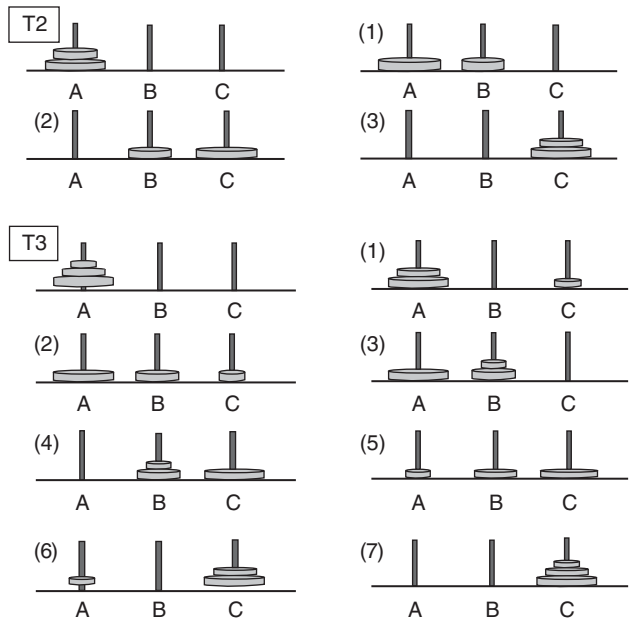
Der rekursive Algorithmus zur Lösung dieses Spiels kann wie folgt zusammengefasst werden:

1. bewege die obersten n – 1 Scheiben von A nach B
2. bewege die n-te Scheibe von A nach C
3. bewege n – 1 Scheiben von B nach C

Der Trick dabei: Sie müssen das rekursive Vorgehen erkennen. Das heißt, jedes Mal, wenn Sie eine Scheibe bewegen, ergibt sich ein neues Problem mit einer kleineren Anzahl an Scheiben, das mit der gleichen Vorgehensweise gelöst werden kann.

Wenn es nur eine Scheibe gibt, müssen Sie nur einen Spielzug machen: Sie legen die Scheibe einfach von A nach C. Wenn es zwei Scheiben gibt, brauchen Sie drei Spielzüge. Die kleine Scheibe legen Sie von A nach B, die große von A nach C, und dann die kleine Scheibe von B nach C – so wie es der Algorithmus vorgibt. T2 stellt diese Spielzugfolge anschaulich dar. Für drei Scheiben brauchen Sie sieben Spielzüge, wie in T3 dargestellt und wie es der Algorithmus beschreibt.

Doch bevor Sie jetzt übermütig werden und es gleich mit, sagen wir mal, zehn Scheiben probieren, sei eins noch gesagt: Die Formel, um herauszufinden, wie viele Spielzüge Sie für wie viele Scheiben brauchen, ist einfach: $2^n - 1$, wobei n die Anzahl der zu bewegenden Scheiben ist. Zum Beispiel: Wenn $n=3$ ist, lautet die Formel: $2^3 - 1 = 8 - 1 = 7$. Zehn Scheiben würden demnach $2^{10} - 1 = 1024 - 1 = 1023$ Spielzüge brauchen!



Die „Türme von Hanoi".

Das rekursive Vorgehen bei den „Türmen von Hanoi" ist tatsächlich sehr einfach: Um n Scheiben von A nach C zu versetzen, wobei B als Zwischenlager dient, verfahren Sie folgendermaßen:

- „wenn n=1 ist" (es gibt nur eine Scheibe) bewegen Sie die Scheibe von A nach C
- „wenn n ≠ 1 ist" bewegen Sie
  - n – 1 Scheiben von A nach B, wobei Sie C als Zwischenlager verwenden

- eine Scheibe von A nach C
- n – 1 Scheiben von B nach C, wobei Sie A als Zwischenlager verwenden

Jetzt müssen Sie nichts weiter tun, als für n – 1 Scheiben genau so zu verfahren, bis nur noch eine Scheibe übrig ist und Sie (abschließend) Schritt 1 (wenn n = 1 ist) ausführen. Für den Fall n = 3, wenn also drei Scheiben auf Stab A liegen, verfahren Sie wie unter „wenn n ≠ 1 ist" beschrieben: Sie lassen zwei Scheiben auf Stab A liegen und versetzen eine auf Stab C. Dann nehmen die nächste Scheibe von Stab A, setzen Sie auf B, legen die Scheibe von C auf B und haben damit n – 1 = 2 Scheiben von A zu B gesetzt. Nehmen Sie nun die Scheibe von A und legen sie auf C (dies ist oben der zweite Schritt). Danach nehmen Sie die oberste Scheibe von B und legen Sie auf A, die verbliebene Scheibe von B legen Sie auf C und haben damit n – 1 Scheiben von B nach C bewegt. Abschließend legen Sie die letzte Scheibe von A auf C (wie unter „wenn n = 1 ist" beschrieben). Fertig!

Ihre Vorgehensweise ist rekursiv, da durch die immer gleiche Zugfolge eine Wiederholung entsteht, solange, bis mehrere aufgetürmte Scheiben auf eine einzige Scheibe reduziert sind und die Wiederholung mit dem letzten Zug dieser einen Scheibe endet. Die Lösung ist in der Box 8.1 anschaulich dargestellt.

Der „Zauberwürfel" (auch Rubiks Würfel genannt) ist ein weiteres Knobelspiel, das über rekursiv definierte Zugfolgen gelöst werden kann. Sobald Sie dahintergekommen sind, wie diese aussehen, ist dieses scheinbar vertrackte Spiel kinderleicht zu lösen.

Doch wenn Sie nun denken, Rekursionen beträfen nur irgendwelche dämlichen Spiele, dürften Sie erstaunt sein zu erfahren, dass Sie sich selbst tagtäglich damit befassen, nämlich jedes Mal, wenn Sie sprechen oder andere sprechen hören und verstehen. Alle natürlichen Sprachen sind das Produkt rekursiver Regelsysteme, um Sätze zu generieren und sie sinnhaft verstehen zu können. In der Tat ist dies einer der machtvollsten Aspekte jeder natürlichen Sprache: Er ermöglicht uns, mithilfe einiger weniger simpler Regeln eine unendliche Vielzahl von Sätzen zu erzeugen. Überlegen Sie doch einmal: Sie haben die Sätze, die hier gedruckt sind, nie zuvor gesehen, und doch können Sie sie unschwer verstehen. Und jeden Tag sprechen Sie Hunderte von einmaligen Sätzen, die Sie nie zuvor so gesagt haben. Und das tun Sie mühelos, oft sogar, während Sie nebenbei etwas ganz anderes tun.

Um zu sehen, wie das funktioniert, wollen wir die folgenden sehr einfachen rekursiven „Formulierungs"-Regeln im Englischen betrachten. Wenn Sie diese einfachen Regeln befolgen, können Sie grammatisch korrekte, englische Sätze in beliebiger Länge bilden. Die Begriffe in Klammern sind optional, alle anderen sind obligatorisch. Die Pfeile (→) bedeuten „ist zusammengesetzt aus".

Satz → Nominalphrase + Verbphrase
Nominalphrase → [Artikel] + [Adjektiv] + Nomen + [Relativsatz]
Verbalphrase → Verb + [Adverb] + [Nominalphrase]
Relativsatz → Relativpronomen + Verb + Nominalphrase

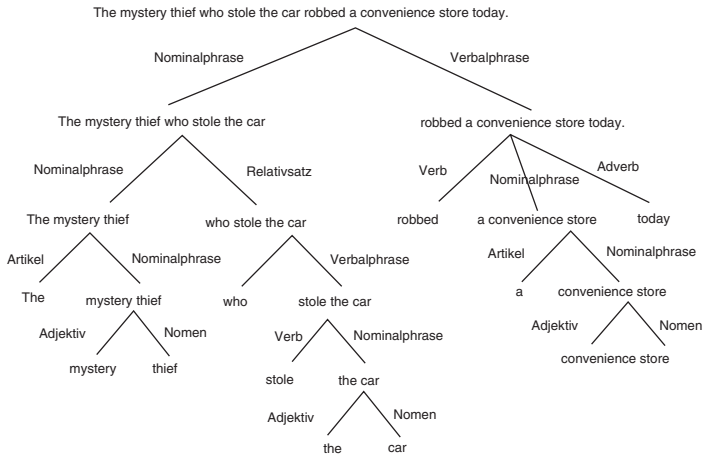


Abb. 8.1 Ein Ableitungsbaum für einen komplexen englischen Satz, konstruiert nach vier einfachen Regeln.

Beachten Sie, dass die Regeln festlegen, wie größere Einheiten (Sätze, Nominalphrasen, Verbalphrasen und Relativsätze) aus kleineren Einheiten (Nomen, Artikel, Adjektive, Adverbien, Pronomen) zu konstruieren sind, und dass sie sich in regelhafter Weise aufeinander beziehen. Nach diesen Regeln können wir den folgenden Satz generieren:

> *The mystery thief who stole the car robbed a convenience store today.*

„*The mystery thief*" ist die Nominalphrase (Artikel, Adjektiv, Nomen); „*robbed a store*" ist die Verbalphrase (Verb, Nominalphrase); „*who stole a car*" ist der Relativsatz (Relativpronomen, Verb, Nominalphrase): Abbildung 8.1 stellt den Ableitungsbaum (in Anlehnung an das Englische auch

als Parse-Baum bezeichnet) für diesen Satz dar und zeigt zudem auf, welche Regeln welchen Teilsatz generieren.

Wir könnten nun beliebig viele Satzteile gemäß dieser rekursiven Regeln hinzufügen, um am Ende Sätze wie vielleicht diesen zu haben:

> *The mystery thief who stole the car that belonged to the French ambassador who was visiting his uncle robbed a convenience store today.*

Wenn Probleme nicht klar definiert sind

Wenn das Leben weitestgehend aus klar definierten Problemen bestünde, die in Algorithmen gepackt ganz leicht zu lösen wären, wären wir wohl die glücklichsten Menschen unter dieser Sonne. Doch leider sind die Probleme, die unser Leben bestimmen, oft recht unklar definiert – und das heißt, *sie lassen klar formulierte Zielzustände oder klar definierte Lösungswege vermissen*. Wie spart man genügend Geld, um im Alter unbeschwert leben zu können? Wie lassen sich Kriege vermeiden? Wie kriege ich diese Frau oder jenen Mann dazu, dass sie/er mit mir ausgeht? Es gibt Algorithmen, die Ihnen sagen, wie Sie das beste Rührei zaubern, aber wie zaubern Sie ein ganz neuartiges Eiergericht, das den Gaumen Ihres Frühstücksgastes verwöhnt? Als englischer Muttersprachler beherrschen Sie die Algorithmen zur Konstruktion englischer Sätze aus dem Effeff, doch was genau sagen Sie, um jemanden von Ihren Ansichten zu überzeugen? Es gibt keinen einvernehmlich vereinbarten

Lösungsweg, um irgendeines dieser Ziele zu erreichen. Es kann viele solcher Wege geben, oder auch keinen.

Unklar definierte Probleme eignen sich nicht für klare Lösungsverfahren. Aus diesem Grund verwenden wir häufig Heuristiken, mit denen wir sie zu lösen versuchen. Heuristiken sind „Faustregeln" oder erfahrungsbasierte Strategien, die die Wahrscheinlichkeit erhöhen, eine erfolgreiche Lösung zu finden. Sie garantieren weder Lösungserfolg, noch garantieren sie korrekte Lösungen. Manchmal verwenden wir Heuristiken auch dann, wenn ein Algorithmus zur Lösung eines Problems verfügbar ist. Wie lautet die Lösung für das Problem „96 × 58"? Wenn Sie einen Taschenrechner parat haben, können Sie dieses Mittel nutzen, um zu einer Lösung zu gelangen. Und wenn nicht? Nun, dann könnten Sie sich mühsam durch den Multiplikationsalgorithmus rechnen, um irgendwann die Lösung zu haben. Was aber, wenn sie versuchen herauszufinden, ob Sie sich 96 m^2 Teppich zu 58 Dollar pro m^2 leisten können? Eine Heuristik würde hier genauso gut funktionieren: 96 gerundet auf 100, macht zwei Nullen an die 58, und prompt hat man die Antwort – 5800 Dollar. Einfache Heuristiken wie diese nehmen ein geistig anspruchsvolles Problem, um es auf ein einfaches Problem herunterzubrechen, das schnell und mit wenig geistiger Anstrengung lösbar ist.

Wie dieses Beispiel zeigt, weisen Heuristiken üblicherweise zwei Merkmale auf, die sie enorm nützlich und praktisch machen: Sie erfordern meist weniger Mühe und weniger Zeit, sie zu implementieren. In einer perfekten Welt hätten wir unbegrenzt Zeit und Verarbeitungsressourcen verfügbar, um Lösungen zu finden. In der realen Welt jedoch, haben wir das oft nicht. Wir müssen Lösungen in

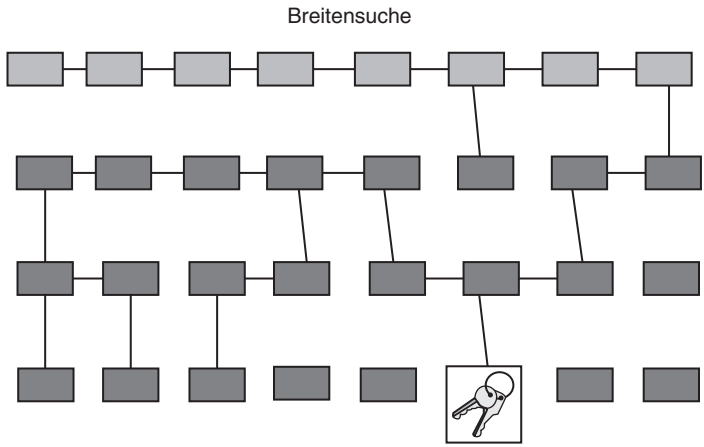


Abb. 8.2 Bei einer Breitensuche werden zuerst alle Möglichkeiten auf einer gegebenen Ebene durchsucht, bevor man sich auf die nächste Ebene begibt.

Realzeit und unter einem angemessenen Suchaufwand finden. Heuristiken sind daher nicht selten unsere Freunde.

Ich will noch einmal das oben erwähnte Beispiel mit den verlorenen Schlüsseln bemühen, um zwischen Algorithmen und Heuristiken zu unterscheiden. Nehmen wir einmal an, Sie machen sich fertig, um aus dem Haus zur Arbeit zu gehen, und können Ihre Schlüssel partout nicht finden. Was nun? Natürlich beginnen Sie, danach zu suchen. Wenn Sie dieses Problem „algorithmisch" angehen, suchen Sie vielleicht systematisch das ganze Haus ab, vom Speicher über jedes Zimmer auf jedem Stockwerk bis hinunter in den Keller.

Wenn Sie jedes Zimmer auf jedem Stockwerk durchsuchen, bevor Sie die Treppe hinab zur nächsten Etage steigen, machen Sie eine sogenannte *Breitensuche*, wie in Abb. 8.2 grafisch dargestellt.

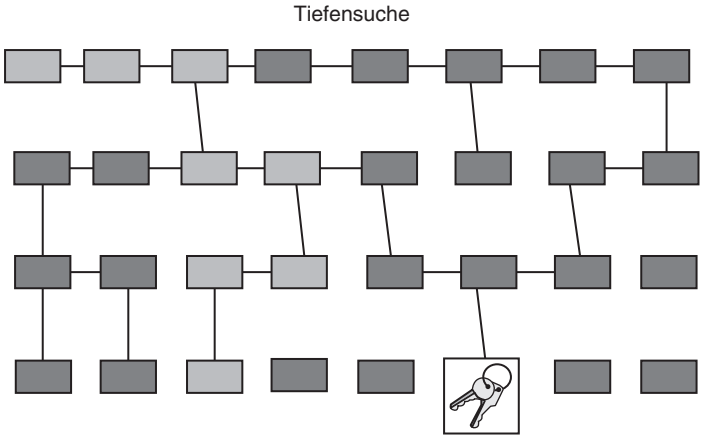


Abb. 8.3 Bei einer Tiefensuche werden zuerst alle Möglichkeiten, die mit jeweils tieferliegenden Ebenen verbunden sind, durchsucht. Danach begibt sich der Suchende wieder zurück auf die oberste Ebene und beginnt die Suche über einen neuen Pfad von oben nach unten.

Wenn Sie aber von jedem Zimmer aus über eine Treppe direkt in das jeweils darunterliegende Zimmer gelangen und Sie jedes Zimmer entlang dieser Treppen von oben nach unten durchsuchen, bevor Sie wieder hinauf in den Speicher zurückkehren, um auch die übrigen Zimmer auf diese Weise zu durchforsten, machen Sie eine sogenannte *Tiefensuche*, wie in Abb. 8.3 dargestellt.

Einem diesem Suchverfahren zu folgen, garantiert, dass Sie erstens mit der Suche irgendwann durch sind und zweitens Ihre Schlüssel finden (sofern sich diese tatsächlich im Haus befinden). Aber dies wäre sehr zeitintensiv.

Wie wäre es, wenn Sie stattdessen eine Heuristik verwendeten? Sie beginnen Ihre Suche an der Stelle, von der Sie

sicher sind, dass Sie die Schlüssel dort zuletzt in der Hand hatten, und folgen dann dem Weg, den Sie von dort aus durch das Haus genommen haben. Sie erinnern sich, dass Sie vergangene Nacht die Haustür aufgeschlossen haben und zum Telefon gerannt sind, das gerade klingelte. Also gehen Sie den Weg von der Haustür zum Telefon ab. Nichts gefunden? Wohin gingen Sie als nächstes? Ins Schlafzimmer. Dann suchen Sie dort. Und so weiter. Im Gegensatz zur systematischen Suche durch jedes Zimmer im ganzen Haus garantiert diese Heuristik nicht, dass Sie die Schlüssel finden werden, selbst wenn sich diese im Haus befinden, aber sie benötigt weniger Zeit und weniger Mühe. Außerdem weist sie erfahrungsgemäß überaus hohe Erfolgsraten auf. (Vielen Dank an Agatha Christies Hercule Poirot dafür, dass er diese Heuristik so schön beschreibt.)

Egal, welches Problem Sie zu lösen versuchen, formal können wir es beschreiben als eine (heuristische) Pfadsuche im sogenannten Problemraum. Der Problemraum ist ein Möglichkeitsraum, der Ausgangszustand, Zielzustand sowie alle möglichen Zwischenzustände umfasst, die bei der Lösung eines Problems auftreten können, egal, ob diese Zustände zielführend sind oder nicht.

Für unser Beispiel der Schlüsselsuche bezieht er sich demnach auf jeden möglichen Raum, den Sie durchsuchen können, sowie die Wege (Pfade), die Sie von diesem Raum aus gehen können. Geht es um die Lösung komplexer Probleme, kann der Problemraum so groß sein, dass die Probleme unlösbar werden. Da der Problemraum eine enorm hohe Anzahl von möglichen Zuständen annehmen kann, versucht der Problemlöser, den Suchraum einzuschränken, und so ist der tatsächliche Lösungspfad, den er tatsächlich

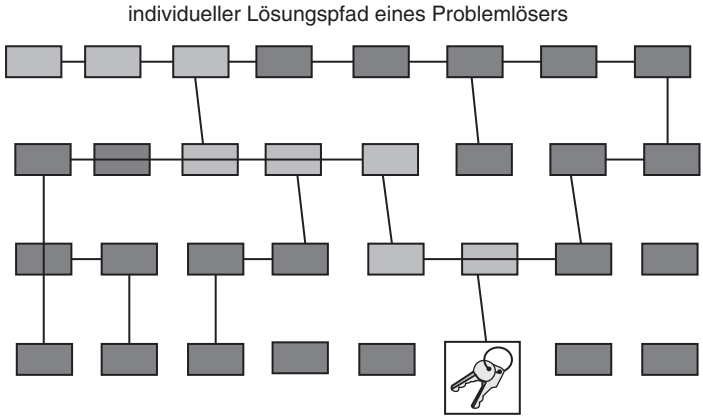


Abb. 8.4 Ein Lösungspfad aus der Menge vieler möglicher Pfade.

geht, für gewöhnlich einer aus der Menge vieler möglichen Pfade, wie in Abb. 8.4 dargestellt.

Die beste Lösung ist natürlich die, die am wenigsten Schritte braucht, um zum Ziel zu kommen.

Es gibt zwei weitere Aspekte der Suche (Problemlösungsprozesse), die hier erwähnt werden sollen. Im eben beschriebenen Szenario haben Sie die Suche mit dem aktuellen Ausgangszustand begonnen (Schlüssel weg!) und Schritt für Schritt das Haus durchsucht, bis das Ziel erreicht war (Schlüssel gefunden!). Diese Art der Suche, die vom Anfang zum Ziel führt, bezeichnet man als Vorwärtssuche (oder Vorwärtsverkettung, *forward chaining*). Aber nehmen wir an, ich hätte Ihnen einen Grundriss Ihres Hauses gezeigt, auf dem ich den Ort, an dem sich die Schlüssel befinden, genau gekennzeichnet hätte und Sie wären ge-

danklich schon auf dem Pfad zurück vom Ziel (Schlüsselfundort) hin zu der Stelle, an der Sie sich jetzt gerade befinden (momentaner Standort). Diese Art der Suche, die vom Ziel zum Anfang führt, bezeichnet man als Rückwärtssuche (oder eine Rückwärtsverkettung, *backward chaining*). Nicht selten bedienen wir uns beider Suchmethoden, wenn wir knifflige Rätsel oder komplexe Probleme zu lösen versuchen. Wir beginnen entweder am Anfang und suchen einen Pfad zum Ziel oder wir beginnen vom Ziel aus und suchen einen Weg zurück zum Anfang. Man findet diese Methoden zum Beispiel auch bei Vernehmungstaktiken oder wenn es darum geht, aus Fakten Schlüsse zu ziehen. Oder nehmen wir einmal an, Sie sind Arzt und suchen nach der richtigen Diagnose für einen Patienten. Um sich zu einer Diagnose vorzuarbeiten, können Sie ihm eine Reihe von Fragen stellen, wie zum Beispiel „Fühlen Sie sich matt?“ oder „Wann haben die Symptome angefangen?“. Sie können aber auch mit einer ersten, unverbindlichen Diagnose beginnen („Hmm, im Moment geht eine schwere Grippe um; vielleicht ist das ja das Problem“), um von dieser Zieldiagnose ausgehend dann die entsprechend zugehörigen Fragen zu stellen („Haben Sie Bauchschmerzen?“ oder „Haben Sie Durchfall?“). Und auch hier können Sie entweder algorithmisch vorgehen und systematisch sämtliche Symptome in jeder Phase des Krankheitsverlaufs durchgehen, oder Sie gehen heuristisch vor, fragen einzelne Symptome ab und stellen mögliche Diagnosen, je nachdem, wie der Patient auf Ihre Fragen reagiert. Anderes Beispiel: Wenn Sie wissen, dass Ihre Tochter am heutigen Tag ein Vorstellungsgespräch sowie eine Prüfung hat, beginnen Sie mit dieser

Information, um Rückschlüsse darauf ziehen, mit welcher Stimmung sie nach Hause kommen wird („Wenn beides gut läuft, wird sie überglücklich sein"; „Wenn nur eins gut läuft, wird sie wie immer sein"; „Wenn beides schief geht, wird sie am Boden zerstört sein"). Sie könnten aber auch abwarten, mit welcher Laune sie nach Hause kommt, um daraus rückzuschließen, wie sie hier wie dort abgeschnitten hat („Sie ist überglücklich, es muss also beides gut gelaufen sein"; „Sie ist wie immer, eins von beiden wird also geklappt haben"; Sie ist am Boden zerstört, es muss also beides schief gegangen sein").

Andere Heuristiken sind von unterschiedlichem Erfolg gezeichnet. So etwa, wenn Sie beschließen, Ihrer Freundin das gleiche Geburtstagsgeschenk zu machen wie im vorigen Jahr, da es ihr damals ja so gut gefallen hat. Oder etwa wenn Sie schätzen sollen, wie viele Männer und Frauen auf der Party letzte Woche waren und dafür nur die Männer und Frauen einrechnen, an deren Gesichter Sie sich erinnern. (Was an diesen Heuristiken verkehrt ist, können Sie sich selbst überlegen.)

Wer suchet, der findet!

Nach Duncker kommen Denkprozesse in Gang, wenn wir mit unserer Ausgangssituation unzufrieden sind und wir keine unmittelbare Idee haben, wie wir das erwünschte Ziel erreichen können. Ziel dieser Denkprozesse ist es, herauszufinden, welche Mittel (oder Handlungen) wir einsetzen können, um unserem Ziel näherzukommen. Der Prozess

des Problemlösens kann daher als Differenzreduktion beschrieben werden – als die Suche nach Handlungen (Operatoren), welche die Differenz zwischen dem gegebenen und dem Zielzustand (maximal) reduzieren. Diese Heuristik kann als Mittel-Ziel-Analyse bezeichnet werden oder auch als heuristische Lösungsstrategie:

1. Zielzustand analysieren
2. Ausgangszustand analysieren
3. Differenzen zwischen Ausgangs- und Zielzustand spezifizieren
4. Spezifizierte Differenzen sukzessiv reduzieren durch
   - direkte Wege und Möglichkeiten, die
   - erreichbare Zwischenzustände, sogenannte Teilziele, generieren

Mit Duncker fand das Phänomen menschlicher Problemlösungsprozesse Einzug in die wissenschaftlichen Labors. Er legte seinen Probanden eine Reihe von unklar definierten Aufgaben vor, die es zu lösen galt, und bat sie, dabei „laut zu denken". Ein inzwischen berühmtes Beispiel solcher Aufgabenstellungen ist das Röntgenproblem: Ein Patient hat einen Tumor im mittleren Bauchraum, der den Patienten töten wird, wenn er nicht entfernt wird. Doch er ist inoperabel aufgrund lebenswichtiger Organe in seiner Nachbarschaft. (Man bedenke, es war das Jahr 1945.) Der Tumor kann zwar durch Strahlung zerstört werden, aber die dazu benötigte Strahlendosis würde mithin das gesamte gesunde Gewebe um den Tumor herum zerstören und den Patienten töten. Wie kann der Tumor zerstört werden,

ohne das gesunde Gewebe zu belasten oder den Patienten zu gefährden? Die Denkprotokolle der Teilnehmer zeigten eine systematische – und oft geistreiche – Suche nach Wegen, den Tumor zu entfernen oder zu schrumpfen, und das gesunde Gewebe möglichst auszusparen. Eine kleine Gruppe von Teilnehmern kam auf eine besonders brillante Lösung, auf die sogenannte Konvergenzlösung. (Worin diese besteht, dürfen Sie gerne versuchen herauszufinden, ehe ich später noch genau darauf eingehen werde.)

Die Mittel-Ziel-Analyse ist keine Problemlösungsstrategie, die sich auf uns Menschen beschränkt. Es ist vielmehr eine natürliche Strategie, die auch im Verhalten anderer Säugetiere in natürlichen Lebensräumen beobachtet werden kann. Die Komplexität des Suchprozesses variiert mit den kognitiven Fähigkeiten der einzelnen Arten. Der Orang-Utan beispielsweise ist ein hochintelligentes Lebewesen, der oft nachahmt, was er sieht, insbesondere, wenn ihm gefällt, was er sieht. Die frei lebenden Orang-Utans auf Borneo sind berühmt berüchtigt dafür, Kanus zu klauen, um darin die unzähligen Flüsse entlangzufahren und sich auf der Suche nach Nahrung auch öfter mal an den Vorräten in Camp Leakey, eine Forschungsstation des World Wildlife Fund (WWF), gütlich zu tun. Eine Orang-Utan-Dame namens Supinah fand besonders Gefallen am Wäschewaschen, das sie sich von den Mitarbeitern der Station abgeschaut hatte, die ihre Wäsche auf einem Waschfloß entlang des Ufers wuschen. Doch die hatten Angst vor ihr und stellten einen Wärter an, um sie fern zu halten, was Supinah nicht im Mindesten abschrecken konnte. Was los war,

wenn Supinah der Belegschaft auf dem Waschfloß einen Besuch abstattete, beschreiben Byrne & Russon (1998):

> „Um am Wärter vorbeizukommen, musste Supinah nur um ihn herum durch das Wasser zum hinteren Ende des Docks kommen, wo sie knietief im Wasser stand und lauerte; sie wusste ganz genau, dass dort, unter dem Dock, ein Einbaumkanu festgebunden war. Für Supinah kein Problem […] prompt machte sie es los und zog es heraus […] richtete es am Dock entlang in Richtung Ziel aus, und steuerte dann flugs darauf zu. Der Rest war ein Klacks – nur noch ein kleiner Sprung mitten hinein in die kreischende Menge auf dem Waschfloß, die bereitwillig Seife und Wäsche fallen ließ und sich ins Wasser rettete. Und Supinah machte sich sogleich ans Werk […] und wusch die Wäsche […]"

Auffallend ist, dass Supinah nicht nur nach Mitteln und Wegen suchte, um den Unterschied zwischen Ausgangszustand und ihrem gewünschten Zielzustand (Wäsche waschen) zu reduzieren, sondern sie generierte auf ihrem Weg zum Ziel auch eine Reihe von Teilzielen wie das Kanu loszumachen und es in die richtige Richtung auf das Ziel hin auszurichten.

Probleme, die, um sie lösen können, zunächst in Teilziele zerlegt werden müssen, sind meist deshalb so schwierig, weil sie das Kurzzeitgedächtnis strapazieren. Man muss einen Überblick über sämtliche dieser Teilziele haben – die erledigten abhaken und die noch nicht erledigten im Auge behalten. Was auch der Grund dafür ist, dass man (zum Beispiel) einen ganzen Berg Wäsche auch ohne Waschmittel waschen kann, dann nämlich, wenn das Kurzzeitgedächtnis das Teilziel „Waschmittel hinzugeben" auf dem Weg zum Ziel verloren hat.

Künstliche Intelligenz: Maschinen, die denken

Nach Dunckers wegweisenden Erkenntnissen über Problemlösungsstrategien möchte man meinen, seine Beiträge hätten unzählige fruchtbare Ansätze im Bereich der psychologischen Forschung hervorgebracht. Aber das ist nicht der Fall. Zu Dunckers aktiver Forscherzeit war die Psychologie (insbesondere in den USA) sehr stark im dem Behaviorismus verhaftet. Ziel des Behaviorismus war es, die Gesetze zu erforschen, die das menschliche Verhalten formen. Forschungen auf dem psychologischen Fachgebiet der klassischen Konditionierung zeigen, dass unser Nervensystem automatisch komplexe Bedingtheiten zwischen einzelnen Ereignissen erlernt, sodass ein Ereignis, das durch einen bestimmten Reiz ausgelöst wird, vorhergesagt werden kann (beispielsweise bedeutet ein Klingeln im Restaurant, dass das Essen gleich serviert wird). Forschungen auf dem Gebiet der instrumentellen Konditionierung zeigen, dass eine bestimmte Verhaltensweise mit einer bestimmten Konsequenz verknüpft sein kann. Das wohl wichtigste Merkmal der behavioristischen Forschung war, dass sie Spekulationen über die inneren Vorgänge im Kopf (in der Blackbox) zu vermeiden suchte und sich strikt auf beobachtbares Verhalten konzentrierte. Der menschliche Geist blieb außen vor, denn innere (mentale) Zustände waren *nicht beobachtbar* – noch nicht. Doch das änderte sich mit der Erfindung des Computers, als plötzlich hoch komplizierte, symbolmanipulierte Rechenmaschinen Probleme lösen konnten und dabei auch eine anschauliche Spur ihres Lösungswegs

hinterließen. Künstliche Intelligenz steckte, wie sich herausstellte, voller innerer Zustände, die ohne Weiteres als Repräsentationen, Suchalgorithmen und Symbolmanipulationen interpretiert werden konnten. Und wenn eine Maschine einen inneren Zustand haben konnte, warum nicht auch der Mensch?

Vor allem die beiden Computerwissenschaftler Allen Newell und Herbert Simon entdeckten Dunckers Arbeiten über Problemlösungsprozesse völlig neu und fragten sich, ob sich Heuristiken wie die Mittel-Ziel-Analyse nicht automatisieren ließen. Könnte man einen Computer nicht so programmieren, dass er Probleme auf dieselbe Weise löst, wie wir Menschen es tun – dass er also in der Lage ist, Differenzen zwischen Ist- und Sollzustand zu reduzieren? Klare Antwort: Ja. 1963 veröffentlichten die beiden Forscher einen Aufsatz, in dem sie die erste erfolgreiche Entwicklung ihrer Programme der künstlichen Intelligenz beschrieben, die logische Theoreme beweisen und logische Probleme lösen konnten. Ihre Programme nannten sie *Logic Theorist* und *GPS* (nein, nicht für Global Positioning System, sondern General Problem Solver). *Logic Theorist* wurde speziell entwickelt, um logische Beweise auszuführen, *GPS* hingegen hatte weitaus ehrgeizigere Ziele, die darin bestanden, Probleme verschiedenster Art nur unter Anwendung der Mittel-Ziel-Analyse zu lösen, so etwa die *Prädikatenlogik erster Stufe*, kryptoarithmetische Probleme, Kannibalen-Missionar-Probleme, die Türme von Hanoi und mathematische Integrale.

Diese bahnbrechenden Programme ebneten den Weg für die Entwicklung von regelbasierten Produktionssystemen –

von Problemlösungsprogrammen, die aus folgenden Komponenten bestehen:

- Bedingungs-Aktions-Regeln (*condition action rules*): Regeln, die genau beschreiben, welche Handlungen zu ergreifen sind, wenn ein Problem unter bestimmten Bedingungen auftritt.
- Langzeitgedächtnis: der Ort, an dem Bedingungs-Aktions-Regeln gespeichert sind.
- Arbeitsgedächtnis: der Ort, an dem Informationen vorübergehend gespeichert und den Regeln angepasst werden.
- Zielstapel: ein „versteckter" Speicher, in den Zwischenziele verschoben werden, bis sie erreicht sind.
- Konfliktlösungsregeln (*conflict resolution rules*): regelbasierte Programme, die alle erkannten Konflikte aus der Menge der Produktionsregeln beheben.

Beispiel: Regeln für das Autofahren:

1. Wenn der Motor aus ist, starten Sie den Motor.
2. Wenn Sie auf Reserve fahren, tanken Sie nach.
3. Wenn es regnet, schalten Sie den Scheibenwischer an.
4. Wenn Sie mit Ihrem Automatikwagen vorwärts fahren wollen, stellen Sie den Schalthebel auf D für *drive* und fahren langsam an.
5. Wenn Sie mit Ihrem Automatikwagen rückwärts fahren wollen, stellen Sie den Schalthebel auf R für *reverse* und fahren langsam an.

6. Wenn die gewünschte Richtung „vorwärts“ ist, die Straße vor Ihnen aber blockiert ist, schalten Sie in den Rückwärtsgang auf „R“ und fahren langsam an.

Nehmen wir an, Regel 4 und auch Regel 6 treffen auf Ihre aktuelle Situation zu. Beide Regeln verlangen ausgeführt zu werden. Aber welche wenden Sie tatsächlich an? Sie könnten diesen Konflikt nach der folgenden Regel lösen: *Wählen Sie die Regel mit der größten Anzahl an Bedingungen*. Das ist Regel 6. Sie enthält mehr Bedingungen als Regel 4 („die gewünschte Richtung ist vorwärts“ und „die Straße vor Ihnen ist blockiert“).

Und so funktioniert nun ein Produktionssystem:

1. Suche mithilfe eines systematischen Abgleichs alle Produktionsregeln, deren Bedingungsteil mit den Dateneinträgen (*inputs*) verträglich ist.
2. Wenn eine passende Regel gefunden ist, führe die Handlung gemäß der vorgegebenen Regel aus.
3. Wenn mehr als eine Regel passt, wähle eine dieser Regeln aus oder ordne sie nacheinander an, um sie gemäß einer vorgegebenen Konfliktlösungsstrategie auszuführen.

Diese äußerst simple und zugleich äußerst wirksame Methode verfährt nach dem Reduktionsprinzip, das heißt, sie reduziert die Differenzen zwischen Ausgangs- und Zielzuständen und ermöglicht es dadurch, Probleme so zu lösen, wie wir Menschen es tun!

In den folgenden Jahrzehnten wurden unzählige Produktionssysteme geschaffen. Einige wurden als Forschungsinstrumente genutzt, um die Möglichkeiten und Grenzen

menschlicher Problemlösungsprozesse aufzuzeigen. Andere wurden zum Zweck der kommerziellen Anwendung geschrieben. Bei diesen Produktionssystemen handelte es sich üblicherweise um sogenannte *Expertensysteme*, um wissensbasierte Systeme: Programme, die die (kognitive) Leistungsfähigkeit menschlicher Experten reproduzierten.

Für die Konzeption solcher Programme untersuchen die Programmierer das Fachwissen der Experten in einem einzelnen Bereich sehr genau, um dieses Wissen dann in einem automatisierten Produktionssystem zu reproduzieren. Die bekanntesten Beispiele medizinischer Expertensysteme sind:

- MYCIN: ein Expertensystem zur Diagnose und Therapieempfehlung von BlutInfektionen, entwickelt von Edward H. Shortliffe und Bruce G. Buchanan an der Stanford University.
- Abdominal Pain System: ein Expertensystem zur Diagnose von akuten Bauchschmerzen, entwickelt von de-Dombal an der University of Leeds
- Health Evaluation Through Logical Processing (HELP): ein Expertensystem zur stationären Krankenhausbehandlung, entwickelt am LDS Hospital in Salt Lake City, das klinische, administrative und finanzielle Funktionen sowie Funktionen zur Entscheidungsunterstützung bereitstellt.

In jüngerer Zeit ist im Bereich der medizinischen Expertensysteme eine enorme Entwicklung zu verzeichnen. Zu den neu etablierten Systemen gehören Akutbehandlungssysteme, Systeme zur Unterstützung von Entscheidungen, Ge-

sundheitsförderung und Qualitätssicherung, Systeme zur medizinischen Bildgebung oder Arzneimittelverwaltung sowie Laborsysteme.

Die Vorteile dieser Systeme sind zahlreich. Sie machen Nutzern, die keinen direkten Zugang zu Experten haben (Ärzte in ländlichen Gebieten etwa), Expertenwissen verfügbar. Sie können wesentliche Informationen abrufbereit speichern. Sie liefern logisch kohärente Antworten für immer wiederkehrende Aufgaben und arbeiten rund um die Uhr.

Expertensysteme haben aber auch Nachteile. Vor allem den, dass sie den gesunden Menschenverstand, der in manchen Entscheidungsprozessen vonnöten wäre, missen lassen. Zudem können sie sich an wechselnde Umgebungen meist nicht adaptieren, es sei denn die Wissensbasis wird gezielt verändert.

Ein schmerzliches Beispiel dafür war der Einbruch der amerikanischen Aktienmärkte, zu dem es am 6. Mai 2010 an der Börse kam, auch Flash Crash genannt. Um 14:45 Uhr EST, stürzte der Dow Jones blitzartig um 9 % (rund 900 Punkte) ab, um sich dann wie durch ein Wunder binnen weniger Minuten wieder zu erholen. Der Handel war an jenem Tag wegen der Sorge einer europäischen Schuldenkrise turbulent. Nach einer Untersuchung der US-Börsenaufsichtsbehörde (Securities and Exchange Commission, SEC) von 2010, hatte die Finanzfirma Waddell and Reed ein Verkaufsprogramm gestartet, das nach einem Algorithmus operierte, der „ohne Berücksichtigung von Preis oder Verkaufszeitraum“ Verkaufsorders platzierte. Dies verursachte einen „Heiße-Kartoffel-Effekt“, bei dem Hochfrequenzhändler begannen, wie verrückt zu verkau-

fen, was einen wahren *run* auf dem Aktienmarkt auslöste. Die Situation wurde innerhalb weniger Minuten erkannt und korrigiert. Zurück aber blieb ein Misstrauen gegenüber hochriskanten Entscheidungsgeschäften, verbunden mit dem Gefühl, sich nicht allzu leichtfertig auf automatisierte Systeme verlassen zu können.

Wie Experten Probleme lösen

Wer korrigiert die automatisierten Problemlöser? Dieselben Leute, deren Kenntnisse und Fähigkeiten man für ihre Produktion herangezogen hatte: Experten. Was aber macht einen sachverständigen Problemlöser aus? Kurzum: Ein Experte zu sein bedeutet, seine Kenntnisse auf einem bestimmten Gebiet ständig zu erweitern (Expertenwissen ist fachspezifisch) und dieses Wissen effizient zu strukturieren, damit Lösungen rasch abgefragt oder konstruiert werden können. Wie wichtig Letzteres ist, kann gar nicht hoch genug geschätzt werden. Je größer die Wissensbasis, umso mehr Informationen gilt es zu durchsuchen und umso länger dauert die Suche. Meist aber ist die Wissensbasis von Experten so organisiert, dass sie auf problemrelevante (oder lösungsrelevante) Eigenheiten besonders reagiert. Und insofern erfolgt die Lösungsabfrage oder die Lösungskonstruktion blitzschnell.

Betrachten wir als Beispiel einmal das strategische Brettspiel Schach und die Experten auf diesem Gebiet, die Schachgroßmeister. Schach ist ein hoch kompliziertes Spiel, das Strategie und vorausschauendes Denken erfordert. Für jede mögliche Figurenkonstellation gibt es in jeder Phase

des Spiels unzählige Zugvarianten. Einige davon sind strategischer als andere, führen zu einer größeren Kontrolle über das Spielfeld und einer höheren Wahrscheinlichkeit, das Spielziel zu erreichen und den Gegner schachmatt zu setzen. Wer also die eigenen Spielzüge überlegt, muss stets auch im Voraus bedenken, welche möglichen Spielzüge sich aus diesem oder jenem Zug für den Gegner ergeben, welche daraus dann wiederum für einen selbst und so fort. Bereits für sechs vorausschauend gedachte Züge können annähernd 1,8 Milliarden verschiedene Spielverläufe entstehen; Schachgroßmeister allerdings schauen selten weiter voraus als 100 Züge in 15 Minuten (Gobet & Simon 1996). Und dennoch schlagen sie immer wieder automatisierte Schachprogramme, die rechenaufwendige Algorithmen bemühen. Tatsächlich hat ein Schachprogramm einen Schachgroßmeister nur ein einziges Mal geschlagen – und das Ergebnis dieses Turniers ist bis heute umstritten. Am 11. Mai 1997 schlug der Schachcomputer Deep Blue, entwickelt von IBM, den damals amtierenden Schachweltmeister Garri Kasparow in der sechsten Matchpartie (zwei Siege für Deep Blue, ein Sieg für Kasparow, drei Remis) (Newborn 1997). Nach der Niederlage warf Kasparow IBM Betrug vor. Er sprach davon, in manchen Zügen des Computers eine hohe menschliche Intelligenz und Kreativität beobachtet zu haben und vermutete daher, seinem Maschinengegner sei durch menschliche Eingriffe auf die Sprünge geholfen worden, was nach den Turnierregeln klar verboten ist. IBM wies die Vorwürfe zurück. Kasparow forderte Deep Blue darauf zu einem Rematch heraus, das im nationalen Fernsehen übertragen werden sollte, doch IBM weigerte sich und

beschloss, Deep Blue aus dem Verkehr zu ziehen. Wie aber könnten diese „menschlichen Eingriffe" ausgesehen haben?

William Chase und Herbert Simon (1973) untersuchten die mentalen Strukturen von Schachspielern auf verschiedenen Spielniveaus – Novizen, Fortgeschrittene, Experten (Simon war selbst einmal ein Großmeister). Sie konnten keinerlei Unterschiede feststellen, was das Regelwissen betraf, den Intelligenzgrad oder die Gedächtnisleistung (Anzahl der Züge, die im Kurzzeitgedächtnis gespeichert werden, sowie Anzahl der vorausgedachten Züge). Gewaltige Unterschiede allerdings zeigten sich darin, wie die einzelnen Spieler die Brettkonstellationen analysierten – das heißt, wie gut sie strategische Anordnungen aufnahmen und sich daran erinnerten. Zu diesem Ergebnis kamen die beiden Forscher, nachdem sie den Spielern fünf Sekunden lang Diagramme mit vorgegebenen Stellungen gezeigt hatten, die aus einzelnen Partien stammten oder Zufallskonstellationen waren, die sie anschließend aus dem Gedächtnis auf echten Schachbrettern mit echten Schachfiguren nachstellen sollten.

Bei Zufallskonfigurationen (die in echten Spielen nicht auftreten konnten) stellten sie keinerlei Unterschiede zwischen Experten und Anfängern fest. Wenn es sich jedoch um echte Konfigurationen handelte, solche also, die in echten Spielen tatsächlich auftreten konnten, waren die Experten in der Lage, aus dem Gedächtnis in vergleichsweise weniger Zeit mehr Figuren in der richtigen Konstellation auf dem Schachbrett nachzustellen, und sie gingen dabei strategisch vor (sie reproduzierten z. B. eine Konfiguration mit rochiertem König und dann ein Läuferfianchetto). Chase und Simon schlossen daraus, dass das Wissen der Schachex-

perten, sprich die strategischen Muster auf einem Schachbrett, zusammen mit den dazugehörigen optimalen Spielzügen, in ihrer Wissensbasis (ihrem Gedächtnisspeicher) gut und strategisch organisiert ist. Und das erklärt, warum sie in der Lage sind, ein Schachbrett visuell sehr schnell zu erfassen, es in strategische Angriffsmuster zu „zerlegen" (z. B. Damenläufer greift Turm an, der vom Damenbauern bewacht wird) und optimale Spielzüge für ein bestimmtes Muster abzurufen.

In neuerer Zeit haben Reingold et al. (2001) eine anspruchsvollere Studie konzipiert, die sie mit Schachnovizen, Fortgeschrittenen und Großmeistern durchführten. Die Ergebnisse, so stellten sie anschließend fest, deckten sich mit denen von Chase und Simons. Den Probanden wurde für eine Sekunde lang ein Schachbrettdiagramm präsentiert, auf dem dann die Position einer einzigen Figur verändert wurde. Beide Diagramme, das ursprüngliche und das veränderte, wurden den Probanden nun jeweils abwechselnd eine Sekunde lang mit einer Pause von 100 Millisekunden zwischen beiden Präsentationen gezeigt, und zwar so lange, bis der Spieler die Veränderung genau erkennen konnte. Einmal mehr zeigte sich ein Unterschied zwischen den einzelnen Gruppen in Bezug auf die Zufallskonfigurationen. Die Experten erfassten die Veränderung im Schnitt eine Sekunde schneller als die Novizen und Fortgeschrittenen. Darüber hinaus zeigte sich, dass alle Spieler ungefähr zehn Quadrate in einer gegebenen Zufallskonstellation erfassen konnten. Präsentierte man ihnen jedoch Konfigurationen aus echten Spielsituationen, konnten die Experten rund 25 Felder erfassen – mehr als das Doppelte!

Auch was die erfassten Spielfiguren betraf, zeigten sich Unterschiede. Die Experten registrierten strategische Figuren visuell häufiger als nichtstrategische und nahmen auch leere Felder zwischen den Figuren häufiger wahr (d. h. die Felder, auf die man ziehen konnte, um andere Figuren anzugreifen oder zu verteidigen). Für die Experten waren diese leeren Felder wichtig und stellten mögliche Angriffsziele dar. Profifußballer, so fand man heraus, machen es nicht anders: Sie beobachten die Bewegungen der Spieler, die gerade nicht in Ballbesitz sind, denen der Ball aber zugepasst werden könnte. Fußballanfänger dagegen neigen dazu, sich nur auf den Ball zu konzentrieren sowie auf den Spieler, der ihn gerade kontrolliert (Williams et al. 1992 und 1994).

Ähnliche Ergebnisse hatten Studien an Physikern, die ebenfalls sehr unterschiedlich auf physikalische Aufgabenstellungen reagierten, je nachdem, ob sie Berufsanfänger oder Experten waren. In einer klassischen Studienreihe von Chi, Feltovich und Glaser (1981) mit Doktoranden der Physik und berufserfahrenen Physikern ging es darum, etliche physikalische Aufgaben aus Standardlehrwerken zu ordnen und zu lösen. Die Doktoranden verfügten über ausreichende Kenntnisse der Physik, um diese Aufgabe bewältigen zu können. Doch sie näherten sich ihnen auf völlig andere Weise als die berufserfahrenen Physiker. Sie ordneten die Aufgaben nach Ähnlichkeit der Oberflächenmerkmale. Zum Beispiel fassten sie Aufgaben, die sich um schiefe Ebenen drehten, zu einer Gruppe zusammen, und solche, die sich um Federn drehten, zu einer anderen. Die berufserfahrenen Physiker dagegen ordneten die Aufgaben nach den der Lösung der Aufgaben zugrunde liegenden Prinzipien. Es überrascht daher nicht, dass die Experten häufiger und schneller zu einer Lösung fanden, als die Doktoranden.

Diese Ergebnisse deuten darauf hin, dass Experten Aufgaben ganz anders „sehen“ als Neulinge auf einem Gebiet, und das liegt zu einem großen Teil daran, dass ihr effizient organisiertes Gedächtnis die Aufmerksamkeit auf lösungsrelevante Merkmale der Problemsituation lenkt. Weil sie an der richtigen Stelle nach dem Richtigen suchen, kommen sie sehr viel leichter zum Ziel. Und aus eben diesem Grund kann ein Schachgroßmeister gegen viele Anfänger gleichzeitig spielen (und gewinnen), ein Beckham immer zur richtigen Zeit am Ball sein und ein Einstein oder ein Feynman Lösungen für physikalische Aufgaben finden, die anderen unlösbar scheinen.

Dem Denken auf der Spur – Erkenntnis und Genius

Ich habe bislang gezeigt, dass die (kognitiven) Prozesse beim Problemlösen als Suchprozesse analysiert werden können. Aber wie können wir „Erkenntnis“ definieren, diesen scheinbar rätselhaften Prozess des Problemlösens, der uns zu plötzlichen Einsichten und Einfällen führt? „Heureka! Ich hab’s!“, so der begeisterte Ausruf, wenn wir wie aus dem Nichts vom Geistesblitz getroffen werden. Dahinter steckt doch sicherlich etwas ganz anderes, oder nicht? Ja und Nein. „Erkenntnis“ beinhaltet bestimmte Komponenten, die nichts mit „normalen“ Prozessen des Problemlösens zu tun haben – insbesondere das „Aha!“-Erlebnis, das uns eine plötzliche Einsicht in die Lösung eines Problems beschert. Neurowissenschaftler, die sich mit diesem Phänomen eingehend beschäftigen, kommen allerdings zu dem Schluss,

dass der Weg zum Ziel auch hier auf komplexe Suchprozesse zurückzuführen ist.

Während der Gestaltpsychologe Karl Duncker eine präzise Beschreibung der „normalen" Problemlösungsprozesse gibt, verändert ein anderer Gestaltpsychologe namens Wolfgang Köhler (1887–1967) unseren „erkenntnistheoretischen" Blick auf das durch Einsicht gewonnene Wissen. Köhler studierte Psychologie und Philosophie an der Universität zu Berlin und leitete danach die Anthropoidenstation der Preußischen Akademie der Wissenschaften auf Teneriffa, wo er auch während der Zeit des Ersten Weltkriegs blieb.

Weit weg von den Kriegswirren daheim setzte er seine Arbeiten ungestört fort, richtete ein riesiges Freigehege ein und untersuchte neun Schimpansen unterschiedlichen Alters. Der Behaviorismus, das beherrschende Theoriegerüst der damaligen Zeit, beschreibt den Vorgang des Problemlösens als einen mehrstufigen Prozess aus Versuch und Irrtum, der ein Verhalten verstärkt, wenn es zu befriedigenden Ergebnissen führt, und ein Verhalten dämpft, wenn es unbefriedigende Ergebnisse nach sich zieht. Köhlers Schimpansen jedoch legten offenbar eine ganz andere Art des Problemlösens an den Tag: De Lösung für eine Aufgabe schien ihnen urplötzlich zu kommen und nicht indem sie ihr Verhalten durch Versuch und Irrtum schrittweise korrigierten. In einem Versuch beispielsweise hängte Köhler Bananen an ein Seil, und zwar so hoch, dass sie für die Schimpansen außer Reichweite waren. Die Schimpansen konnten springen, wie sie wollten, vergeblich, die Bananen waren einfach nicht zu kriegen. Doch plötzlich begannen sie, Kisten, die in der Nähe lagerten, zu stapeln, bis der Kletterturm hoch

genug war, um die Früchte zu erreichen. Sie kooperierten sogar, um das behelfsmäßige Podest zu stabilisieren. Sultan, Köhlers intelligentester Affe, tat sich sogar besonders hervor, weil er einzelne Stöcke ineinandersteckte und sich auf diese Weise ein Werkzeug baute, das lang genug war, um sich die Banane zu angeln, die vor seinem Käfig baumelte. Zuvor hatte Sultan mit jedem Stock einzeln probiert, an die Banane zu kommen, doch sie waren alle zu kurz. Während er hin und her probierte, steckte er zufällig einen Stock in einen anderen, was die Reichweite des Stocks verlängerte. Ohne zu zögern benutzte er sein neues Werkzeug *sofort* in angemessener Weise, holte sich die Banane und ließ damit erkennen, dass er *unmittelbar* erkannt hatte, welch geniale Zufallsentdeckung er gerade gemacht hatte.

Köhler beschreibt dieses Lösungsverhalten als „Einsicht" (im Sinne einer Problemerkenntnis mit zweckgerichtetem Folgeverhalten). Genauer gesagt, er definierte „Einsicht" als eine *plötzliche Wahrnehmung von Beziehungen zwischen den Elementen einer Problemsituation.* Dieses blitzartige Verständnis führt häufig zu einer Umstrukturierung der Problemsituation. Und genau darin liegt der Schlüssel zur (einsichtsvollen) Lösung: Die meisten kreativen Lösungswege zeichnen sich dadurch aus, dass sie spontane Elemente beinhalten, die zu einer Reorganisation problemrelevanter Gedächtnisinhalte führen, und zwar in einer Art und Weise, wie sie für ein gerade bestehendes Problem atypisch ist.

Doch wie kommen diese plötzlichen Einsichten zustande, und wann? Auf diese Fragen gibt es zwei erklärende Antwortmodelle. Das erste besagt, dass einsichtsvolle Lösungen spontan erfolgen (*special-problem explanation*): Der Problemlöser wechselt urplötzlich von einem Zustand des

völligen Unwissens in einen Zustand des völligen Wissens. Das zweite Erklärungsmodell besagt, dass einsichtsvolle Lösungen das Ergebnis gradueller Prozesse sind, die sich schrittweise vollziehen und den Problemlöser langsam an das Ziel heranführen (*business as usual explanation*).

In den 1980er-Jahren erforschten Janet Metcalfe et al. (1986, 1987) systematisch und mit ziemlich originellen Methoden das „normale" und „einsichtige" Problemlösen. Sie führten Studien mit drei verschiedenen Aufgabentypen durch, die jeweils von allen Teilnehmern gelöst werden mussten:

- inkrementelle Aufgaben (Nicht-Einsichtsaufgaben), die mithilfe von Algorithmen auf Gymnasialniveau (schrittweise) gelöst werden konnten ($3 \times 2 + 2x + 10$),
- inkrementelle Aufgaben, die mithilfe einfacher heuristischer Suchverfahren gelöst werden konnten (Sie haben Behältnisse mit unterschiedlichem Volumen, 4,6 l, 0,4 l, 0,7 l und 0,3 l, sowie einen unbegrenzten Vorrat an Wasser. Aufgabe ist es, ein Wasservolumen von 2,2 l abzumessen. Sie dürfen die Behältnisse nach Belieben füllen und wieder ausleeren und Wasser von einem Krug in einen anderen umschütten; siehe auch Box 8.2, Wasserumfüllproblem) und
- nichtinkrementelle Aufgaben (Einsichtsaufgaben), wie etwa diese: Seerosen vermehren sich in bestimmten Gebieten alle 24 Stunden um das Doppelte. Zu Beginn des Sommers gibt es eine Seerose auf dem See. Es dauert 60 Tage, bis der See komplett mit Seerosen bedeckt ist. Am wievielten Tag ist der See zur Hälfte mit Seerosen bedeckt?

Alle 15 Sekunden wurden die Probanden aufgefordert, auf einer Skala von 1 bis 7 „Kalt-Heiß-Bewertungen" darüber abzugeben, wie nahe sie sich der Lösung wähnten, wobei 1 für kalt stand („Ich bin völlig ahnungslos") und 7 für heiß („Ich bin sicher, die Antwort zu haben").

Die Ergebnisse für die inkrementellen Aufgaben (Nicht-Einsichtsaufgaben) schienen die *special-process explanation* zu untermauern: Die Einschätzung nahmen im Verlauf langsam aber stetig von 1 bis 7 zu, da sich die Probanden zunehmend der Lösung der Nicht-Einsichtsaufgaben näherten. Für die Einsichtsaufgaben (die nichtinkrementellen Aufgaben) jedoch gaben die Probanden durchgängig und immer wieder Bewertungen zwischen 1 und 2 ab, bis zu dem Punkt, da sie wie aus dem Nichts heraus auf die Lösung kamen und die volle Bewertungszahl von 7 gaben. Mit anderen Worten, sie wechselten spontan von einem Zustand der völligen Ahnungslosigkeit in einen Zustand der völligen Einsicht (gemäß der *special-problem explanation*).

Bowers et al. gingen 1990 in eine etwas andere Richtung, indem sie ihren Probanden sogenannte CRAProbleme (*compound remote associates*) präsentierten, Aufgaben, die durch eine Abfolge bewusster und miteinander assoziativ verknüpfter Ideen gelöst werden können. Die Aufgaben bestanden zum Beispiel aus Worttriaden, Gruppen mit jeweils drei Wörtern, für die es jeweils ein viertes Wort zu finden galt, das am Anfang oder am Beginn aller drei Wörter der Gruppe sinnhaft angefügt werden konnte, sodass sich daraus ein neues Wort ergab. Hier einige Beispiele:

A *french*, *car*, *shoe*
B *boot*, *summer*, *ground*
C *table*, *wall*, *dog*

Für die ersten beiden Worttriaden gibt es tatsächlich Lösungen (*horn* bzw. *camp*), für die dritte jedoch nicht. Die Wörter haben begrifflich nichts gemeinsam. Den Probanden wurden immer zwei Triaden gleichzeitig präsentiert, wobei eine lösbar war und eine nicht (zum Beispiel A und C, siehe oben). Jede Präsentation dauerte nur wenige Sekunden. In dieser Zeit mussten die Probanden erstens die richtige Lösung nennen (sofern sie eine finden konnten), zweitens eine Entscheidung darüber abgeben, welche der beiden Triaden sie für lösbar halten (auch wenn sie weder für die eine noch für die andere eine Lösung gefunden hatten), und drittens anhand einer Skala von 0 bis 2 bewerten, wie sicher sie sich mit ihrer jeweils abgegebenen Entscheidung waren (0 = „keinesfalls sicher", 2 = „sehr sicher").

Interessanterweise waren die Teilnehmer in der Lage, ziemlich korrekt einzuschätzen, für welche Dreiergruppe es tatsächlich eine Lösung gab und für welche nicht, was die Logik der *special-problem explanation* allerdings ins Wackeln bringt: Wenn nämlich einsichtsvolle Lösungen aufgrund spontaner und unverrückbarer Erkenntnisse erfolgen, dann wären die abgegebenen Entscheidungen der Probanden darüber, welche Dreiergruppen lösbar sind und welche nicht, nichts weiter als ein Zufallsergebnis. Obendrein hätten sie sich in einem „ahnungslosen" Zustand befinden müssen, sofern sie für keine der beiden Triaden eine Lösung gefunden hatten. Auf die Frage, welche von beiden Triaden sie für lösbar halten, hätten sie also nur raten können und hätten damit wohl in 50 % der Fälle richtig gelegen. Doch das war nicht der Fall. Im Schnitt lagen sie mit ihrer Einschätzung in rund 60 % der Fälle korrekt, sprich, sie hatten die CRA-Aufgabe als lösbar erkannt. Und dieses Ergebnis

liegt statistisch höher als ein Zufallsergebnis. Gut möglich also, dass im Denken der Probanden kleine, entscheidende Teilschritte zur Lösungsfindung stattgefunden haben, nur eben außerhalb der bewussten Wahrnehmung.

Um mehr Klarheit über die neuronalen Abläufe bei der Wahrnehmung zu bekommen, leisten die neurowissenschaftlichen Verfahren der neueren Zeit, insbesondere die funktionelle Magnetresonanztomografie (fMRT), gute Dienste. Diese Technik nutzten auch Bowden et al. (2005). Sie führten Studien durch, im Rahmen derer die Probanden CRA-Aufgaben sowie Nicht-Einsichtsaufgaben lösen sollten, während gleichzeitig eine fMRT durchgeführt wurde, die Bilder ihrer aktivierten Hirnareale lieferte. Mithilfe dieser Technik war es den Forschern möglich, genau zu beobachten, welche Hirnareale während des Vorgangs des Problemlösens aktiv waren. Was sie herausfanden, warf bestehende Vorstellungen über das Erkennen der Lösung eines Problems oder des Lösungswegs für eine praktische Aufgabe völlig über den Haufen.

Für Einsichtsaufgaben verzeichneten die Forscher eine vergleichsweise stärkere Aktivierung der rechten Hirnhälfte. Das eigentlich Interessante aber ist, dass die Aktivität kurz bevor eine Einsichtsaufgabe gelöst wurde, abrupt von der linken zur rechten Hirnhälfte wechselte. Dieser abrupte Wechsel war bei Nicht-Einsichtsaufgaben nicht zu beobachten. So weit, so gut, was scheinbar „plötzliche" Einsichten anbelangt (*special-process explanation*) anbelangt. Doch werfen wir nun einen Blick auf die Ergebnisse, die zeigen, dass Einsicht auch eine *business as usual*-Komponente hat.

Die Prozesse, die der Lösung von Einsichtsaufgaben zugrunde liegen, stellten sich folgendermaßen dar: Zunächst

zeigte sich eine starke Aktivierung der dominanten (linken) Hemisphäre und eine schwache Aktivierung in der nichtdominanten (rechten) Hemisphäre. Dies bedeutet, dass die dominanten Bedeutungsinhalte für die jeweiligen Wörter einer Triade stark aktiviert waren. Das Problem dabei ist, dass diese dominanten semantischen Assoziationen nicht sonderlich hilfreich sind. Die dominante Assoziation für „*french*" (eine Sprache, gesprochen in Frankreich) hat sehr wenig zu tun mit der für „*car*" (Transportfahrzeug für jedermann) oder der für „*shoe*" (Schutzkleidung für die Füße). Aber es gibt auch nichtdominante Bedeutungsinhalte dieser Wörter und die können mit allen drei Wörtern der Triade semantisch sinnhaft assoziiert werden („*french horn*", „*car horn*", „*shoe horn*"). Diese nichtdominanten Bedeutungsinhalte sind ebenfalls aktiviert, wenn auch weniger stark. Binnen weniger Sekunden breitet sich diese schwächere Aktivierung zunehmend aus und wird immer stärker, insbesondere, wenn die dominanten Bedeutungsinhalte schwächer werden. (d. h., wenn man aufhört, unentwegt darüber nachzudenken). Abbildung 8.5 zeigt eine mögliche Darstellung dieser Vorgänge.

Sind die nichtdominanten Bedeutungsinhalte ausreichend aktiviert, „schießen" sie ins Bewusstsein und der Problemlösende erkennt urplötzlich die Lösung. Dies zeigte sich als eine verstärkte Aktivität im sogenannten Alphaband (in einem anderen Hirnareal) ungefähr 1,5 Sekunden vor der Gammawelle, die das spontane Lösen einer Aufgabe begleitete. Die Probanden hatten an der Lösung der Aufgabe also tatsächlich außerhalb ihrer bewussten Wahrnehmung gearbeitet. Und dieser Lösung wurden sie erst gewahr, als sie ihnen schlagartig ins Bewusstsein schoss, obwohl sie auf

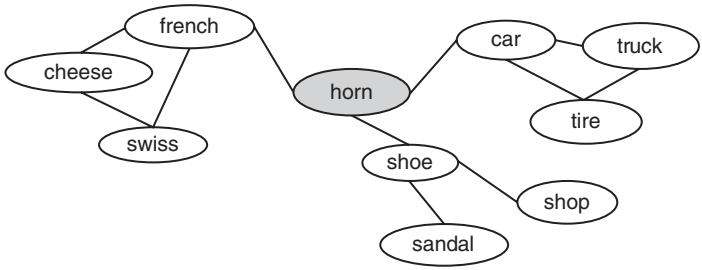


Abb. 8.5 Automatischer Suchvorgang, der häufig unbewusst abläuft.

der unbewussten Ebene die ganze Zeit ihr Gedächtnis nach lösungsrelevanten Inhalten durchforstet haben. Einsicht scheint folglich eine besondere Mischung zu sein aus einer schrittweisen Suche durch den Problemraum (*business as usual*) und einem plötzlichen Aha-Erlebnis (*special-process*).

Wenn Einsicht ein einfacher, mehrstufiger Prozess im Unterbewusstsein ist, warum tritt er dann so selten zutage? Im Schnitt produzieren wir einsichtsvolle Lösungen in nur 25 Prozent der Fälle. Die Antwort scheint darin zu liegen, dass Einsicht (wie auch andere Aspekte der kognitiven Wahrnehmung) den sogenannten Rahmungseffekten (*framing effects*) unterliegt: Die Art und Weise, wie ein Problem präsentiert (beschrieben) wird, beeinflusst unser lösungsorientiertes Denken. Und dies wiederum begrenzt uns vor allem dann, wenn die Lösung einer Aufgabe kreative Ansätze erfordert.

Der sogenannte Einstellungseffekt ist dafür ein perfektes Beispiel. Damit bezeichnet man die mechanische Anwendung von zuvor angewendeten Lösungsmustern auf neue Aufgaben, die aber zu negativen Lösungserfolgen führen,

weil eingefahrene Denkvorgänge das kreative Problemlösen beeinträchtigen. Der Einstellungseffekt erklärt auch, warum es manchmal vorkommt, dass ein Novize plötzlich die Lösung für eine Aufgabe findet, an der Experten zuvor gescheitert sind. Der Effekt wurde in verschiedenen Kontexten nachgewiesen. Bekannt sind vor allem die Studien des Gestaltpsychologen Abraham Luchins (1942, 1959). Luchins stellte mehrere Probanden vor das „Wasserumfüllproblem". Aufgabe war es, aus einer unbegrenzten Menge Wasser und unter Zuhilfenahme dreier Behältnisse mit jeweils unterschiedlichem Volumen eine bestimmte Menge Wasser abzumessen und dabei möglichst effizient vorzugehen. Einige Beispiele sind der Box 8.2 aufgeführt. Sie dürfen sich gerne selbst einmal an dieser Aufgabe versuchen. Gehen Sie die vorgegebenen Beispiele einfach der Reihe nach durch.

Box 8.2 „Wasserumfüllproblem" nach Luchins

Problem	Volumen Krug A	Volumen Krug B	Volumen Krug C	gewünschte Zielmenge
1	21	127	3	100
2	14	163	25	99
3	18	43	10	3
4	9	42	6	21
5	20	59	4	31
6	23	49	3	20
7	15	39	3	18
8	28	76	3	25
9	18	48	4	22
10	18	48	4	22

Lösungen:
Alle Aufgaben bis auf Aufgabe 8 können auf dieselbe Weise gelöst werden: B – 2C – A. Die Aufgaben 1 bis 5 lassen sich mit dieser Formel am einfachsten lösen. Doch für die Aufgaben 7 und 9 gibt es auch eine einfachere Lösung (A + C), ebenso wie für die Aufgaben 6 und 10 (A – C). Aufgabe 8 kann mithilfe von B – 2C – A nicht gelöst werden, mithilfe von A – C jedoch sehr wohl.

Sind die Probanden auf diese einfacheren Formeln gekommen? Nein! 83 % von Luchins Versuchsteilnehmern gingen für die Aufgaben 6 und 7 nach der Formel B – 2C – A vor. Sie hatten sich in ihrer Routine festgefahren und konnten die einfachere Lösung schlicht nicht sehen. Die Versagerquote für Aufgabe 8 lag bei 64 %, da die Probanden versuchten, die Formel anzuwenden, die sie bislang erfolgreich angewendet hatten. Doch als sie damit nicht weiterkamen, waren sie in ihrem produktiven Denken behindert und konnten die einfachere Formel auch nicht finden. Für die Aufgaben 9 und 10 benutzten 79 % der Probanden die Formel B – 2C – A, obgleich diese ebenfalls mit einer einfacheren Formel zu lösen gewesen wären.

Als Luchins den Probanden nur die letzten fünf Aufgaben stellte, benutzte weniger als 1 % die Formel B – 2C – A und nur 5 % scheiterten an Aufgabe 8.

Und als er ihnen den warnenden Hinweis gab, ab Aufgabe 5 nicht mit Scheuklappen zu denken, waren über 50 % imstande, die einfachere Lösung für die restlichen Aufgaben zu finden.

Das Interessante dabei: Die ersten fünf Aufgaben lassen sich allesamt auf dieselbe Weise, sprich mit der gleichen (etwas komplizierten) Formel lösen. Für die übrigen Aufgaben gibt es auch eine einfachere Lösung (Formel). Doch bis zu Aufgabe 6 hatten sich die Probanden in „ihrem Fahrwasser fest-

gefahren“ und versuchten, auch die übrigen Aufgaben wie gehabt zu lösen. Dies bedeutete, dass Aufgabe 8 überhaupt nicht gelöst werden konnte, obgleich es für die Aufgaben 6 bis 10 eine einfache Lösung gibt. Mit ihrer bis dahin erfolgreichen, wenn auch komplizierten Lösungsstrategie, die sie weiterhin auf die übrigen Probleme zu übertragen suchten, standen sich die Probanden mit einem Mal selbst im Weg – und konnten für Aufgabe 8 gar keine Lösung finden.

Sehr anschaulich dargestellt ist dieses Suchverfahren auch in Abb. 8.3. Angenommen, Sie machen eine Tiefensuche und beschreiten zunächst einen Suchpfad, der Sie vollständig in die Tiefe führt, weit weg vom Ziel und ohne Möglichkeit, es zu erreichen. Dann ist es am besten, Sie gehen wieder ganz nach oben und beginnen Ihre Suche von neuem. Oder noch besser, gönnen Sie sich eine kleine Pause und beginnen Sie Ihre Suche danach.

In der Fachwelt spricht man bisweilen von einem Inkubationseffekt, wenn eine Aufgabe nach erfolglosen Lösungsversuchen eine Weile unterbrochen wird und danach schnell und erfolgreich gelöst werden kann. Betrachten wir folgende Aufgabe, die von Silveira entwickelt wurde (1971):

> „Sie bekommen vier Halsketten mit jeweils drei Kettengliedern, die Sie zu einer vollständigen, geschlossenen Halskette zusammenfügen sollen. Das Öffnen eines Kettengliedes kostet zwei Cent, das Schließen eines Kettengliedes kostet drei Cent. Zu Beginn des Experiments sind alle Kettenglieder geschlossen. Fügen Sie alle zwölf Kettenglieder so zusammen, dass die ganze Kette insgesamt nicht mehr als 15 Cent kostet.“

Silveira stellte für dieses Experiment zwei Versuchsgruppen zusammen. Die Bearbeitungszeit für beide Gruppen betrug eine halbe Stunde, wobei eine der beiden Gruppe eine halbstündige Pause machen und sich mit anderen Dingen beschäftigen musste. In der Gruppe ohne Pause fanden 55 % der Probanden zu einer Lösung, in der Gruppe mit Pause waren es 64 %. In einer dritten Gruppe, der ebenfalls eine allerdings fünfstündige Pause zugestanden wurde, lösten sogar 85 % der Probanden die Aufgabe!

Mit Duncker fand das Phänomen menschlicher Problemlösungsprozesse Einzug in die wissenschaftlichen Labors. Er legte seinen Probanden eine Reihe von unklar definierten Aufgaben vor, die es zu lösen galt, und bat sie, dabei „laut zu denken". Ein inzwischen berühmtes Beispiel solcher Aufgabenstellungen ist das Röntgenproblem: Ein Patient hat einen Tumor im mittleren Bauchraum, der den Patienten töten wird, wenn der Tumor nicht entfernt wird. Doch die Geschwulst ist aufgrund lebenswichtiger Organe in seiner Nachbarschaft inoperabel.

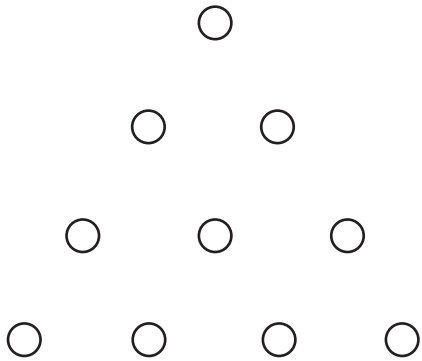
Kommen wir an dieser Stelle noch einmal zurück auf Duncker und das Röntgenproblem. Wie Sie sich vielleicht erinnern, lautete das Problem, dass ein Patient einen Tumor im mittleren Bauchraum hat. Dieser Tumor wird den Patienten töten, wenn er nicht entfernt wird. Doch die Geschwulst ist aufgrund lebenswichtiger Organe in seiner Nachbarschaft inoperabel. Haben Sie eine Lösung gefunden? Wenn nicht, denken Sie doch einmal über folgende Aufgabe nach: Ein Physiker führt ein Experiment durch und verwendet dafür eine speziell konstruierte, sehr teure und sehr zerbrechliche Glühbirne. Der Glühfaden in der

Birne reißt, kann aber mit der Hitze einer Laserdiode repariert werden. Laserdioden können Licht in verschiedenen Wellenlängen aussenden, die verschiedene Temperaturen haben. Das Problem: Die Temperatur, die nötig ist, um den Glühfaden wieder zusammenzufügen, würde das Glas der Glühbirne schmelzen. Der Physiker grübelt eine Weile über dieses Problem und besorgt sich aus benachbarten Labors etliche Laserdioden. Diese positioniert er rings um die Glühbirne herum und richtet sie so aus, dass die Lichtstrahlen allesamt auf den Glühfaden zulaufen. Dann stellt er die Laserdioden auf eine niedrige Wärmestufe ein und schaltet sie an. Sobald die Strahlen konvergieren, summieren sich die Temperaturen und erreichen rasch die hohe Temperatur, die vonnöten ist, um die Enden des Glühfadens zusammenzufügen. Wissen Sie jetzt, wie Sie Dunckers Tumorpatienten retten können? Wenn nicht, lesen Sie das nächste Kapitel, in dem ich auf Analogiebildungen eingehe, ein weiteres Verfahren, um Probleme zu lösen.

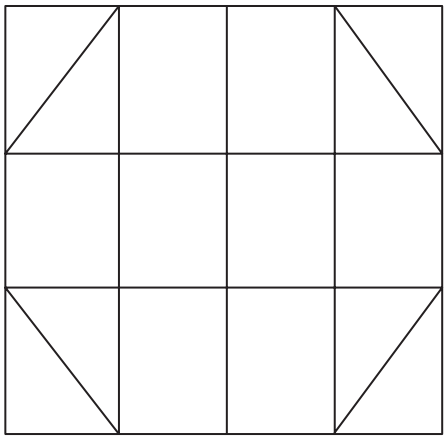
Und wenn Sie sich an weiteren Einsichtsproblemen versuchen wollen, schauen Sie sich die von Metcalfe und Wiebe (1987) in Box 8.3 an.

Box 8.3 Einsichtsprobleme – Bitte lösen!

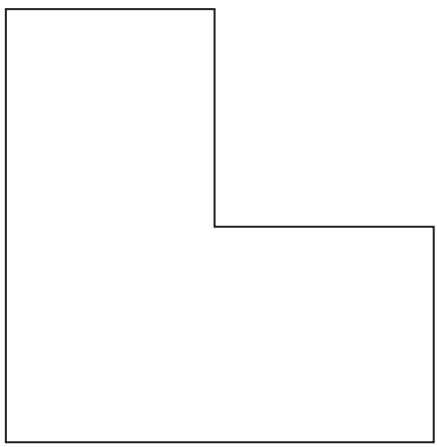
1. Schneiden Sie in eine Karteikarte ein Loch, das groß genug ist, um Ihren Kopf hindurchzustecken. Schaffen Sie es? Nur zu, es gibt einen Weg.
2. Das Dreieck, das Sie hier sehen, zeigt mit der Spitze nach unten. Verschieben Sie drei Kreise so, dass das Dreieck mit der Spitze nach oben zeigt.



3. Ein Mann kauft ein Pferd für 60 Dollar und verkauft es für 70 Dollar. Dann kauft er es für 80 Dollar zurück und verkauft es wieder für 90 Dollar. Wie viel Gewinn hat er aus seinem Pferdegeschäft gemacht?
4. Eine Frau hat vier Halsketten mit jeweils drei Kettengliedern, die Sie zu einer ganzen Halskette zusammenfügen will. Das Öffnen eines Kettengliedes kostet zwei Cent, das Schließen eines Kettengliedes kostet drei Cent. Sie hat aber nur 15 Cent, die sie dafür ausgeben kann. Wie stellt sie es an?
5. Ein Landschaftsgärtner erhält den Auftrag, vier Bäume so zu pflanzen, dass jeder genau im gleichen Abstand zum anderen steht. Wie würden Sie es machen?
6. Eine kleine Schale Öl und eine kleine Schale Essig stehen nebeneinander. Sie nehmen einen Esslöffel Öl und rühren ihn in den Essig. Dann nehmen Sie einen Esslöffel dieser Mischung und rühren ihn wieder in die Schale mit dem Öl. Welche der beiden Schalen ist stärker „kontaminiert"?
7. Zeichnen Sie die Figur, die Sie hier sehen, nach, ohne dabei den Stift abzusetzen, das Papier zu falten oder bereits gezeichnete Linien noch einmal abzufahren. Wie schaffen Sie das?



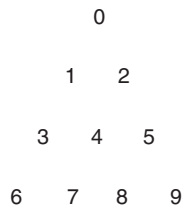
8. Sie sollen 27 Tiere auf vier Gehege verteilen, wobei sich in jedem Gehege nur eine ungerade Zahl an Tieren befinden darf. Wie machen Sie das?
9. Teilen Sie die Figur, die Sie hier sehen, in vier gleiche Teile, und zwar so, dass alle Teile dieselbe Größe und dieselbe Form haben. Wie gehen Sie vor?



10. Sie sollen zehn Münzen so anordnen, dass sich am Ende fünf Reihen (Linien) mit jeweils vier Münzen ergeben. Wie machen Sie das?

Lösungen:

1. Die Antwort darauf finden Sie unter: http://pbskids.org/zoom/activities/games/stepthroughhole.html
2. Ich habe die Kreise nummeriert, um die Lösung besser nachvollziehbar zu machen. Kreis 6 und Kreis 9 werden in die (noch) zweite Reihe verschoben, der Kreis mit der 0 kommt nach unten.



3. 20 Dollar (–60+70–80+90).
4. Und so geht's: Zu Beginn haben wir vier Ketten mit jeweils drei Kettengliedern.

000 000 000 000

Um eine vollständige, geschlossene Kette zu bekommen, müssen wir nur drei Kettenglieder öffnen.

– – – 000 000 000

Das kostet uns sechs Cent, zwei Cent pro geöffnetem Kettenglied.

Die drei geöffneten Kettenglieder fügen wir nun jeweils an eines der anderen drei Kettenteile an:

000 – 000 – 000 –

Dann schließen wir sie, wobei wir die zwei Enden aneinanderfügen. Das kostet uns neun Cent, und zwar drei Cent, um jedes Glied zu schließen. 6 + 9 = 15 Cent.

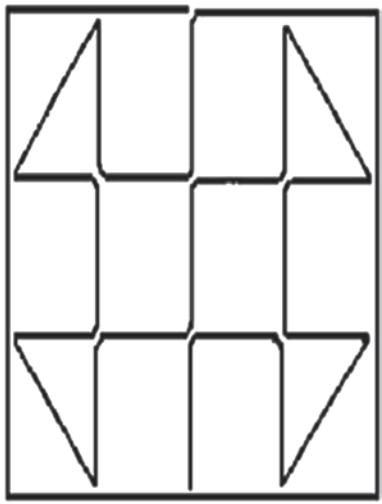
5. Pflanzen Sie die Bäume auf einen Hügel. Drei im gleichen Abstand um den Hügel herum und einen auf die Kuppe. Der Abstand zwischen den Bäumen am Fuße des Hügels zum Baum auf der Kuppe sollte jeweils der gleiche sein wie der zwischen den Bäumen am Boden.
6. Beide sind gleichermaßen „kontaminiert". Und zwar aus folgendem Grund:

   Zu Beginn: Schale A enthält 2 Esslöffel (EL) Öl;
   Schale B enthält 2 EL Essig

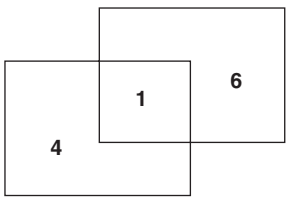
   Schritt 1: Schale A enthält 1 EL Öl; Schale B enthält 2 EL Essig + 1 EL Öl; ein EL dieser Mischung enthält also 2 / 3 Essig und 1 / 3 Öl.

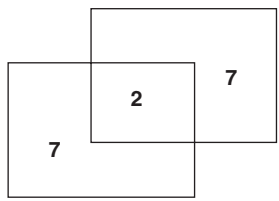
   Schritt 2: Schale A enthält 1 1 / 3 EL Öl und 2 / 3 EL Essig; Schale B enthält 1 1 / 3 EL Essig und 2 / 3 EL Öl.

7. Beginnen Sie an der Stelle, wo an der oberen Linie die kleine Lücke ist: Folgen Sie der Linie nach links und dann immer weiter rings um die Figur herum, bis Sie wieder an der Lücke angelangt sind. Von dort aus folgen Sie der Linie nach und wechseln Sie an jeder Lücke die Richtung.

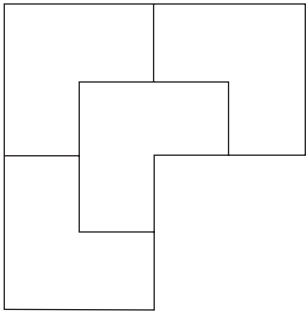


8. Der Trick dabei ist, dass sich die Gehege überlappen müssen, sodass sich einige der Tiere genau genommen in zwei Gehegen befinden. Hier eine Lösung:

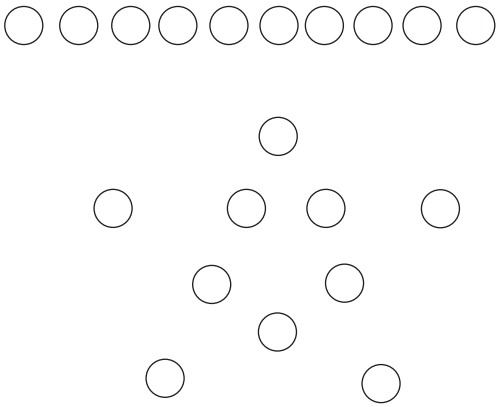




9. Eine Möglichkeit, die Figur in vier gleiche Teile zu teilen, ist diese:



10. Eine Möglichkeit, die zehn Münzen in fünf Reihen (Linien) mit jeweils vier Münzen anzuordnen, besteht darin, sie wie unten gezeigt horizontal und schräg zu positionieren.



9

Analogieschlüsse

Das ist wie jenes

In den Annalen der Finanzgeschichte sticht das Jahr 2008 hervor wie eine Vogelspinne auf Weißbrot. 2008 nämlich war das Jahr, in dem die Bankenindustrie in ihre schlimmste Krise seit der Großen Depression schlitterte. Nie dagewesene Steigerungen der Immobilienpreise während des Jahrzehnts zuvor hatten Banker verleitet, immer riskantere Hypothekendarlehen auszugeben. Als die Immobilienblase platzte, platzten auch die Hypothekenportfolios. Die US-amerikanische Großbank Lehman Brothers ging Pleite. Andere, wie Merrill Lnych oder AIG, entgingen einer Pleite nur um Haaresbreite, ehe die US-Regierung einschritt, um die Banken zu retten, die als *too big to fail*, als „zu groß?, um pleite zu gehen", erachtet wurden.

Dies war ein enormes und potenziell unpopuläres Unterfangen, weshalb sich der Präsident des Federal Reserve Board, Ben Bernanke (Notenbankchef), bemüßigt fühlte, der Nation in einer einstündigen Fernsehansprache zu erklären, warum es unumgänglich sei, das kriselnde Bankensystem zu retten. Und das tat er, indem er eine Analogie heranzog. Stellen Sie sich vor, so Bernanke, Ihr Nachbar im Haus nebenan wäre ein verantwortungsloser Mensch, der im Bett raucht und damit sein Haus in Brand steckt.

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5_9,

Sollten Sie die Feuerwehr rufen oder sich einfach umdrehen, weil Sie sich sagen, dass er mit den Konsequenzen seines Handelns selbst fertig werden muss? Und was ist, wenn Ihr Haus (und auch alle anderen in der Nachbarschaft) aus Holz gebaut sind? Wir sind uns alle darin einig, so fuhr Bernanke fort, dass es unter diesen Umständen zuallererst darum gehen müsse, den Brand zu löschen. Und erst dann können wir überlegen, wer Schuld hat, wer strafrechtlich belangt werden muss, ob die Brandschutzverordnung erneuert werden muss oder Brandmelder aufgestellt werden sollen.

Doch es soll hier nicht um Verdienste oder Mängel der staatlichen Bankenrettungspakete gehen. Es geht vielmehr um die Macht der Analogie als kraftvolles Mittel, um zu erklären und zu überzeugen. Eine kraftvolle Analogie kann eine größere Wirkung entfalten als alle Theoriestunden, Schaubilder oder Dokumentationen zusammen. Bernankes Analogie vom „Brand im Nachbarhaus" war genial, denn der durchschnittliche Fernsehzuschauer konnte dieses Bild ohne Weiteres verstehen (der Brand kann auch auf mein Haus übergreifen und es zerstören) und auf die Finanzkrise übertragen (wenn diese Großbanken Pleite gehen, dann ist auch mein Geld und mein Kreditgeber futsch). Doch auch eine noch so starke Analogie kann nach hinten losgehen. Warum? Weil sie für gewöhnlich stark überfrachtet ist und ebenso gut auch andere Interpretationen stützen kann.

So wollte auch Bernankes Analogie nicht allen schmecken. Für viele Finanzanalysten und Politologen war sie ein gefundenes Fressen, das sie scharf durch die Mangel drehten. Wenn Sie bei Google „Bernankes Brandanalogie" eingeben, erhalten sie rund 181 000 Treffer. Auf der Website

des Centre for Research on Globalization finden Sie einen Artikel von Wirtschaftsprofessor Michael Hudson, der die Schwachstellen der Analogie besonders deutlich macht (http://informationclearinghouse.info/article22229.htm):

Was ist falsch an dieser Analogie? Zunächst einmal befinden sich Banken nicht in derselben Nachbarschaft, in der die meisten Menschen leben. Sie sind vielmehr die Burg auf dem Hügel, die Herren über die Stadt zur ihren Füßen. Warum sie nicht abbrennen lassen und den Hügel der Natur zurückgeben, anstatt sie zu retten, damit die ganze Stadt hinaufblickt zu einem Tempel des Geldes, der die Bürger in Schulden geknechtet hält? Mehr als falsch ist auch die Analogie mit der US-Politik. Notenbanken und Finanzpräsidenten „löschen keinen Brand". Sie bemächtigen sich der Häuser, die vom Brand unversehrt geblieben sind, setzen Eigentümer und Bewohner vor die Tür, um die Immobilie auf jene Übeltäter zu übertragen, die „ihr eigenes Haus niedergebrannt haben". Die Regierung spielt nicht die Feuerwehr. „Den Brand löschen" heißt nichts anderes als eine Entschuldung der Wirtschaft, als die Schulden zu löschen, die diese „Wirtschaft niederbrennen".

Warum war das Bild vom „brennenden Haus" für die Fernsehzuschauer so leicht zu verstehen und für die Kritiker so leicht es in der Luft zu zerreißen? Die Kognitionswissenschaftler Keith Holyoak und Paul Thagard (1989) behaupten, dass das „analoge Denken ganz einfach dem Denken eines normalen menschlichen Wesens entspricht". Douglas Hofstadter, ebenfalls Kognitionswissenschaftler und Pulitzerpreisträger beschreibt die Analogie an sich als „das Herz der Erkenntnis". Entwicklungspsychologen haben mehrfach festgestellt, dass Kinder diese Fähigkeit in jungen Jah-

ren von selbst ausbilden, ohne direktes Zutun von Eltern oder Lehrern. Analoges Denken scheint im menschlichen Geist von Natur aus angelegt zu sein. Was genau es damit auf sich hat, werden wir in diesem Kapitel erfahren.

Analogie in der Theorie

Unter einer Analogie versteht man eine relationale (oder auch strukturelle) Ähnlichkeit. Analogien sind ein zentraler Punkt menschlicher Erkenntnis und setzen voraus, dass Objekte, Ereignisse oder Sachverhalte isomorph sind, dass sie also gewisse (strukturelle) Merkmale gemeinsam haben. Mit anderen Worten, zwei Objekte oder Ereignisse sind nicht analog, weil sie sich auf dieselben Dinge beziehen, sondern weil die Beziehungsverhältnisse zwischen diesen Dingen untereinander gleich sind. Ein brennendes Haus hat rein physisch nichts gemeinsam mit einer angeschlagenen Bank. Auf einer tieferen Wahrnehmungsebene jedoch beschreiben beide ähnliche Beziehungsverhältnisse: Ein brennendes Haus ist vom physischen Kollaps bedroht, eine angeschlagene Bank vom finanziellen Kollaps. Und beide, und das ist wohl der sehr viel wichtigere Aspekt, sind auf einschreitende Maßnahmen angewiesen, wenn der Kollaps verhindert werden soll.

Wir könnten also sagen, dass die Oberflächenmerkmale einer Analogiesituation variieren können, die Beziehungen unter ihnen aber die gleichen sind. Gemeint sind die Beziehungen zwischen den beiden Analogiebereichen „Basis" (*base*) und „Ziel" (*target*). Das Ziel einer Analogiebildung ist das Objekt oder das Ereignis, das wir zu verstehen ver-

suchen (wie beispielsweise die Bankenrettung). Die Basis einer Analogiebildung ist das Objekt oder das Ereignis, mit dem das Ziel verglichen wird. Wenn wir Analogieschlüsse vornehmen, bilden wir relationale Merkmale des Ziels auf die Basis ab, sodass beide miteinander korrespondieren. Dies führt dazu, dass wir das Ziel in der gleichen Weise verstehen wie die Basis. Auf Bernankes Analogie angewendet: „Haus" korrespondiert mit „Bank", „brennen" korrespondiert mit „Wertverlust" und „Nachbarschaft" korrespondiert mit „deine Bank und dein Darlehenskonto". Die Lösung für das brennende Haus besteht darin, Löschwasser hineinzupumpen. Der Analogieschluss, will heißen die Lösung für die angeschlagene Bank, besteht darin, Geld in die Bankenindustrie zu stecken. Das Wasser stoppt das Feuer und das Bankenrettungsgeld stoppt den Wertverlust der Anlagekonten.

Wie dieses Beispiel zeigt, besteht die Macht der Analogien darin, dass sie es möglich machen, Informationen von der Basis auf das Ziel zu übertragen. Sie sind ein gutes Hilfsmittel zum besseren Verständnis und verhelfen uns nicht selten zu einer möglichen Lösung für ein Problem. Streifen wir die Oberflächenmerkmale einer Analogiesituation einmal ab, blicken wir darunter auf ein abstraktes Lösungsschema, das man für diese Art von Problemsituation in etwa so formulieren könnte: Wann immer einer lebenswichtigen Ressource die Gefahr einer vollständigen Zerstörung droht, gilt es nachzupumpen, um die Zerstörung zu stoppen. Man beachte, dass dieses Schema mit keiner Silbe Banken, Häuser, Brand, Wertverlust, Wasser oder Geld erwähnt. Dennoch lässt es sich auf beide Szenarien (Haus-

brand und Bankenpleite) anwenden, wie auch auf all jene Szenarien, die diese relationalen Strukturen teilen.

Sich auf Analogien zu verlassen, um Probleme zu verstehen und zu lösen, ist aber insofern problematisch, da Analogien ein bisschen so sind wie die sprichwörtlichen Elefanten im Porzellanladen. Werden sie nicht im Zaum gehalten, können sie eine Menge Schaden anrichten, was in einem umso größeren Chaos enden kann. Wie wir im vorangegangenen Kapitel gesehen haben, fällt die Mechanisierung des Denkens nicht selten dem Einstellungseffekt anheim. In früheren Erfahrungsmustern zu verharren, kann das Auffinden einer neuen, einfacheren oder überhaupt einer Lösung eines bestehenden Problems verhindern. Warum? Ganz einfach: Wenn die Ähnlichkeit zwischen einem bestehenden und einem vergangenen Problem eine gedankliche Analogiebildung anstößt, verleitet dies automatisch dazu, dieselbe Lösungsstrategie auf das bestehende Problem anzuwenden –, auch wenn sie nicht funktioniert oder man sich die Sache dadurch schwieriger macht, als es sein müsste. Bis heute hat die Forschung keine Patentrezepte (Algorithmen) finden können, um Analogieschlüsse zu unterbinden. Stattdessen hat sie erkannt, dass das bewusste Aktivieren von Heuristiken und Denkmustern uns sehr häufig aus diesen Schwierigkeiten heraushelfen kann.

Betrachten wir ein einfaches Beispiel, das uns zurück auf die Schulbank in den Physikunterricht versetzt. Auf dem Stundenplan steht das Rutherford'sche Atommodell. Die Lehrerin beginnt mit einer kleinen Einführung und sagt: „Ein Atom ist wie das Sonnensystem." Das Ziel (*target*) ist das Atom (das, was sie uns Schülern versucht, begreiflich zu machen). Die Basis (*base*) ist das Sonnensystem (von dem

sie annimmt, dass wir es schon kennen). Dann fährt sie fort: „Der Atomkern ist wie die Sonne und die Elektronen des Atoms sind wie die Planeten." Dieser Satz verleitet, das Basismerkmal „Sonne" auf das Zielmerkmal „Atomkern" und das Basismerkmal „Planeten" auf das Zielmerkmal „Elektronen" abzubilden. Sie erklärt: „Die Elektronen kreisen um den Atomkern, so wie die Planeten um die Sonne kreisen." So weit, so gut. Welche Schlüsse können wir anhand dieser Analogie nun ziehen? Wie wäre es mit diesem? „Die Sonne ist größer als ein Planet, also ist der Atomkern größer als ein Elektron." Okay, das kann man so sagen. Und wie wäre es mit diesem? „Die Sonne ist heißer als ein Planet, also ist der Atomkern heißer als ein Elektron." Nein, das kann man so nicht stehen lassen. Aber warum nicht? Welcher Grundsatz (oder welche Grundsätze) erlauben die ersten beiden Analogieschlüsse, schließen den dritten aber aus?

Dedre Gentner entwickelte 1983 die sogenannte *structure mapping theory*, eine Strukturtransfertheorie des analogen Denkens, die in ihrem Kern zwei einschränkende Prinzipien beschreibt. Das erste ist das *relation-mapping principle*: Es besagt, dass Analogien mit ähnlichen Relationen gegenüber jenen mit ähnlichen Attributen bevorzugt werden. („Attribute" ist hier ein anderer Begriff für „Oberflächenmerkmale"). Das zweite ist das *systematicity principle*: Es besagt, dass Analogien, die kohärente Beziehungssysteme erlauben (die also stärker mit anderen Relationen vernetzt sind und sozusagen ein System von Relationen darstellen), eher aufeinander abgebildet werden als individuelle Beziehungen. Dies bedeutet, dass wir Prädikate, die stärker mit anderen Relationen vernetzt sind, den Vorzug geben sollten gegenüber

Prädikaten, die einfach nur auf einzelne Attribute bezogen sind. Schauen wir uns einmal an, wie das funktioniert.

Zunächst stellen wir die Information, die wir haben, in Bezug auf relationale Begriffe (Prädikate) und Attribute dar:

Masse (Sonne) Masse (Atomkern)
Masse (Planet) Masse (Elektron)
umkreist (Planet, Sonne) umkreist (Elektron, Atomkern)
größer als (Masse [Sonne], Masse [Planet])
ist heiß (Sonne)

Anhand des *relation-mapping principle* sehen wir, dass dies eine ziemlich gute Analogie ist: Die Attribute können nicht aufeinander abgebildet werden („Sonne" und „Planet" stehen auf der einen Seite, „Atomkern" und „Elektron" auf der anderen), alle relationalen Begriffe jedoch können aufeinander abgebildet werden („Masse" und „umkreist"). Wenden wir nun das *systematicity principle* an: Dieses Prinzip besagt, dass wir Relationen höherer Ordnung bevorzugt aufeinander abbilden sollten. „Ist heiß" ist eine individuelle Relation; sie bezieht sich auf nur ein einziges Attribut und kann andere Relationen nicht vernetzen. Insofern ist es eine schlechte Wahl, von der Basis auf das Ziel zu transferieren. Aber „größer als" vernetzt zwei Relationen, die sowohl im Ziel als auch in der Basis auftauchen. Insofern können wir diese Beziehung problemlos transferieren. Die Analogie „größer als [(Masse [Sonne], Masse [Planet]) kann also sinnvoll angewendet werden.

Die Kognitionspsychologen Keith Holyoak und Paul Thagard (1997) erweiterten Gentners *structure mapping theory* um die *multi-constraint theory*. Sie bemisst die Kohärenz

einer Analogie nach drei Kriterien, die zu einer optimalen Interpretation der Analogie zwischen Basis und Ziel führen: Struktur (strukturelle Konsistenz), Semantik (semantische Ähnlichkeit) und (pragmatischer) Zweck. Im Kern bevorzugt diese Theorie Analogien, die eine strukturelle Kohärenz aufweisen, und betrachtet in Anlehnung an Gentner die relationalen Strukturen zwischen Basis und Ziel (*systematicity principle,* Relationen höherer Ordnung werden bevorzugt aufeinander abgebildet). Sie bemisst zudem auch Analogien, die auf Ähnlichkeiten in der Bedeutung der Relationen und Attribute basieren, auf die eine Analogie Bezug nimmt. Und sie fragt nach dem Einfluss des pragmatischen Zwecks, der die Auswahl der Analogie bestimmt. Diese multiplen Einschränkungen (*multiple constraints*) sorgen dafür, dass Analogien in einer realistisch sinnvollen Weise ausgewählt und abgebildet werden, die es ermöglicht, ein bestimmtes Ziel zu erreichen.

Analogie in der Praxis

Soweit der Stand der Forschung zu optimalen Analogien. Aber wenden wir sie genau so auch an, um Probleme zu lösen, Entscheidungen zu treffen oder Überzeugungsarbeit zu leisten?

Betrachten wir zunächst Anwälte in einem Gerichtssaal. Anwälte müssen Geschworene und Richter überzeugen, indem sie stichhaltige Argumente vortragen, die auf solider rechtlicher Logik aufgebaut sind. Es mag Sie überraschen zu hören, dass Analogien dabei eine wichtige Rolle spielen. Sie werden häufig herangezogen, um Parallelen zwischen ei-

nem noch unentschiedenen und einem bereits entschiedenen Fall zu ziehen. Anwälte und Richter werden das Urteil aus einem vorangegangenen Fall auf einen zu befindenden Fall anwenden, wenn sie zu der zwingenden Überzeugung gelangen, dass es hinreichende Ähnlichkeiten zwischen beiden Fällen gibt. Wenn sich beide Fälle perfekt aufeinander abbilden lassen, spricht man von einer „Punktlandung" der Beweisführung (einem „Volltreffer", wie man im allgemeinen Sprachgebrauch sagt). Da Anwälte und Richter in dieser Art von stichhaltiger Beweisführung besonders geschult sind, können sie sicherlich als Experten des analogen Denkens gelten – und das auf höchstem Niveau.

Ein sehr schönes Beispiel dafür ist der Fall Adams vs. New Jersey Steamboat Company, New York, von 1896. Adams war Passagier eines Dampfschiffs auf einer nächtlichen Fahrt von New York nach Albany. Während er schlief, wurde ihm sein Geld gestohlen, obwohl er Kabinentür und Fenster verschlossen hatte. Der Fall befasste sich mit diesem wesentlichen Punkt: Ist ein Dampfschiff eher ein „schwimmendes Hotel" oder ehe eine „Eisenbahn auf Wasser"? Ähnelt es eher einem „schwimmenden Hotel", wäre der Beklagte (der Schifffahrtbetreiber) haftpflichtig, denn so will es das Gesetz für Gastwirthaftung. Ähnelt es aber eher einer Eisenbahn auf Wasser, dann wäre der Beklagte (der Schifffahrtbetreiber) nicht haftpflichtig, denn Eisenbahngesellschaften übernehmen keine Haftung für das Eigentum ihrer Fahrgäste. Sie sind nur dann haftbar zu machen, sofern ihnen ein grob fahrlässiges Verschulden nachgewiesen werden kann.

Das Gericht entschied, ein Dampfschiff sei einem „schwimmenden Hotel" ähnlicher als einer „Eisenbahn auf

Wasser“ aufgrund folgender Logik: Hotels haben Zimmer für Gäste, Dampfschiffe haben Zimmer für Passagiere. Hotelgäste wie Dampfschiffpassagiere bezahlen für die Nutzung der Zimmer aus jeweils denselben Gründen – um eine private Unterkunft für die Nacht zu haben. Infolgedessen sind beide gleichermaßen dem Risiko ausgesetzt, bestohlen zu werden, was eine „spezielle Beziehung“ erzeugt zwischen „Hotel und Gast“ auf der einen Seite und „Dampfschiff und Passagier“ auf der anderen Seite. Diese „spezielle Beziehung“ bringt eine gewisse Eigenverantwortung für den Hotelgast bzw. den Schiffspassagier mit sich. Die Analogie zwischen den beiden Vergleichsobjekten wurde vom Gericht offenbar als derart stark befunden, dass es in seinem Urteil formuliert: „Die beiden Relationen, wenn auch nicht verhältnisgleich, weisen eine derart enge Analogie zueinander auf, dass dem Haftungsanspruch Genüge geleistet werden muss.“ Und damit hat das Gericht das Gesetz für Gastwirthaftung auf Dampfschiffe transferiert.

Das Gericht prüfte auch die Gegenanalogie, wonach ein Dampfschiff einer „Eisenbahn auf Wasser“ ähnelte, wies sie aber aufgrund eines wesentlichen Unterschieds zurück: Der Dampfschifffahrtsbetreiber übernimmt die vollständige Betreuung seiner Passagiere, indem er ihnen ein privates Zimmer zur ausschließlichen Nutzung bereitstellt. Eisenbahngesellschaften tun dies nicht. Sie stellen nur wenigen Passagieren eine begrenzte Schlafmöglichkeit in Form offener Kojen zur Verfügung.

Wie dieses Beispiel zeigt, orientierte sich das Gericht in seiner Urteilsfindung weitgehend daran, welche Analogie die größte Gesamtähnlichkeit mit dem unentschiedenen Fall aufwies und insoweit am besten darauf passte. Und

ebenso deutlich zeigt sich, dass das Gericht auch zwischen Oberflächengemeinsamkeiten und strukturellen Parallelen differenzierte. Es wurden Präzedenzfälle herangezogen, in denen es um Hotels und Eisenbahnen ging, da eine Überlappung der Oberflächenmerkmale feststellbar war. Aber diese Oberflächenmerkmale waren nicht das entscheidende Kriterium. Vielmehr wurde der Fall aufgrund struktureller Ähnlichkeiten entschieden, nicht aufgrund von Oberflächenähnlichkeiten. Ein privates Zimmer zur ausschließlichen Nutzung bereitzustellen, impliziert eine größere Haftungsübernahme seitens der Betreibergesellschaft als eine offene Koje in einem Schlafwagen zuzuteilen, wo ein Diebstahlrisiko weniger vermeidbar ist. Meist immer aber läuft es darauf hinaus, dass der Anwalt den Fall gewinnt, der die überzeugendsten Analogien vorbringt.

Die Strategie hinter Analogiebildungen lässt sich vielleicht am besten wie folgt zusammenfassen: Wenn Sie nicht sicher sind, wie Sie ein bestimmtes Problem lösen sollen, suchen Sie nach einem ähnlichen Problem, von dem Sie wissen, wie Sie es lösen können und wenden Sie diesen Lösungsweg auf das vorliegende Problem an. Im Einzelnen bedeutet das:

1. Suchen Sie nach einem ähnlichen Problem, von dem Sie wissen, wie Sie es lösen.
2. Finden Sie heraus, welche strukturellen Korrespondenzen es zwischen dem neuen und dem alten Problem gibt.
3. Wenden Sie die Lösung auf das neue Problem an.

Welcher Schritt ist der schwierigste? Ob Sie es glauben oder nicht, es ist Schritt 1. Natürlich greifen wir immer ganz

automatisch auf Probleme oder Fälle zurück, die wir aus der Vergangenheit kennen, um ähnlichkeitsbasierte Vergleiche anzustellen. Doch beziehen sich die Ähnlichkeiten tendenziell auf Oberflächenstrukturen, die nicht sonderlich hilfreich sind, um daraus Lösungen abzuleiten.

Im letzten Kapitel haben wir Dunckers Röntgenproblem betrachtet: Ein Patient hat einen inoperablen Tumor im mittleren Bauchraum. Der Tumor kann zwar durch Strahlung zerstört werden, aber die dazu benötigte Strahlendosis würde mithin das gesamte gesunde Gewebe um den Tumor herum zerstören und den Patienten töten. Sie fanden dieses Problem wahrscheinlich sehr schwierig zu lösen, bis Sie einen Abschnitt weiterlasen. Dort war von einem Physiker zu lesen, der ein Experiment durchführt, im Zuge dessen der Glühfaden in einer Glühbirne reißt. Mit einer Laserdiode könnte er die Enden des Glühfadens wieder zusammenfügen, doch die Temperatur, die dafür nötig ist, würde das dünne Glas der Glühbirne schmelzen. Daraufhin positioniert er mehrere Laserdioden mit einer jeweils niedrigen Temperatur rings um die Glühbirne herum, und zwar so, dass die Lichtstrahlen allesamt auf dem Glühfaden zusammentreffen (konvergieren), wodurch sich die einzelnen Temperaturen summieren und rasch die hohe Temperatur erreichen, die vonnöten ist, um die Ende wieder zu verbinden. Die Ähnlichkeit zwischen Röntgen- und Glühbirnenproblem dürfte Ihnen nicht entgangen sein, weshalb Sie die „Konvergenzlösung" für das Glühbirnenproblem nun auf das Röntgenproblem anwenden: Sie stellen mehrere Röntgenapparate rings um den Patienten auf, die alle jeweils eine niedrige Strahlendosis aussenden und zur gleichen Zeit auf dem Tumor zusammentreffen.

Aber was, wenn ich Ihnen statt der Glühbirnengeschichte die folgende Geschichte erzählt hätte? Glauben Sie, Sie hätten ebenfalls Ähnlichkeiten erkannt?

> Eine Burg, umzogen von einem Burggraben, ist durch mehrere schmale Brücken mit dem Land verbunden. Eine angreifende Armee nimmt die Burg erfolgreich ein, indem nur wenige Soldaten gleichzeitig über jede Brücke vordringen und sternförmig auf die Burg zulaufen.

Die Antwort auf meine obige Frage lautet sehr wahrscheinlich „Nein". Nur 10 % der Probanden lösen das Röntgenproblem mit der „Konvergenzlösung", wenn das Problem allein (also ohne ein Vergleichsproblem) präsentiert wird. Präsentiert man ihnen die Burggeschichte zuerst (deren analoge „Konvergenzlösung" darin besteht, dass, gleichzeitig an der Festung angekommen, die Summe aller Kämpfer ausreicht, um die Festung einzunehmen), verbessert sich die Erfolgsquote zwar, allerdings nicht wesentlich; nur 30 % der Probanden erkennen die Ähnlichkeit zwischen dem Röntgenproblem und der Burggeschichte und wenden die „Konvergenzlösung" an (Gick & Holyoak 1980). Präsentiert man ihnen jedoch die Glühbirnengeschichte zuerst, verbessert sich die Erfolgsquote sehr viel drastischer; 70 % erkennen die Ähnlichkeit zwischen dem Röntgen- und dem Glühbirnenproblem auf Anhieb und wenden die Konvergenzlösung erfolgreich an (Holyoak & Koh 1987).

Warum dieser Unterschied? Weil das Abrufen von Erinnerungen stark von Oberflächenähnlichkeiten gesteuert ist, insbesondere bei Novizen auf einem Gebiet. Trotz der Tatsache, dass das Röntgen-, das Glühbirnen- und das Burg-

problem alle identische Problemstrukturen haben, differiert die Ähnlichkeit ihrer Oberflächenmerkmale enorm. Röntgenstrahlen, so das allgemeine Empfinden, sind Laserstrahlen ähnlicher als Soldaten. Gleichwohl ist empfindliches Gewebe einem zerbrechlichen Glas ähnlicher als schmalen Brücken. Und so kommt es, dass sich die Probanden eher an die Geschichte über Laserdioden oder die Glühbirne erinnern, wenn sie im Anschluss daran von Röntgenstrahlen und empfindliches Gewebe lesen, und nicht an die Geschichte über Soldaten und Brücken.

Der Kognitionspsychologe B. H. Ross (1984) demonstrierte dies in einer Reihe von Studien über die Verwendung von bildhaften Beispielen für das kognitive Lernen. Den Probanden wurden drei mathematische Formeln vorgelegt, die sie lernen sollten (wie etwa Permutations- oder Kombinationsformeln). Zu jeder Formel gab es ein Beispiel in Form einer Textaufgabe aus verschiedenen Themenfeldern (Pizzaservice oder ein Betrunkener, der seine Schlüssel sucht o.ä.). Dann wurden den Teilnehmern ein Test präsentiert, der dieselben Formeln und dieselben Textaufgaben enthielt, nur war jetzt jeder Formel eine andere Textaufgabe zugeordnet als während der Lernphase. Nehmen wir einmal an, als lehrreiches Beispiel für die Permutationsformel diente die Geschichte eines Betrunkenen, der seine Schlüssel sucht, und für die Kombinationsformel diente die Geschichte eines Burschen, der Pizzen ausliefert. In der Testphase wurden diese beiden Geschichten vertauscht, sodass nun umgekehrt die Geschichte des Betrunkenen für die Kombinationsformel stand und die Geschichte über den Pizzalieferanten für die Permutationsformel. Die Probanden waren aufgefordert, „laut zu denken", während sie

die Aufgaben lösten, sodass das Abrufen von Erinnerungen nachvollziehbar blieb. Welche Aspekte der Aufgaben würden die Stichworte geben und als auslösende Reize für das Abrufen der Erinnerungen wirken? Wie würden die Probanden reagieren? Würden sie sagen: „Oh, wieder eine Pizzaaufgabe", wenn sie versuchen, die neue Pizzalieferantengeschichte zu lösen, oder würden sie sagen: „Oh, wieder eine Permutationsaufgabe"? Welcher Aspekt der Aufgabe würde als auslösender Reiz wirken?

Die Antwort war eindeutig: Mehr als 70 % der Erinnerungen waren inhaltlich gesteuert. Und das ist nicht überraschend, denn wenn die Probanden sagen: „Oh, wieder eine Pizzaaufgabe", anstatt: „Oh, wieder eine Permutationsaufgabe", würden sie zu einer falschen Lösung kommen, denn sie würden versuchen, die neue Pizzaaufgabe unter Verwendung der Kombinationsformel zu lösen, die dieser in der Lernphase zugeordnet war.

Warum kann ein Novize analoge Aufgaben anhand von Oberflächenähnlichkeiten abrufen? Weil der Erfolg der Suche nach einer analogen Aufgabe nur so gut sein kann wie die Qualität der „Datenbank" (der Wissensbasis), die er durchsucht. Wenn man etwas Neues lernt, weiß man zunächst nicht, was wirklich wichtig ist zu wissen und was nicht. Oberflächenmerkmale werden geradewegs zusammen mit allen wichtigen Dingen abgespeichert und oft stärker wahrgenommen (da lebendiger oder verständlicher) als Lerninhalte, auf die es eigentlich ankommt. Es gilt also, die angelegte Wissensbasis zu reorganisieren, um von irrelevanten Oberflächendetails zugunsten zugrunde liegender (relationaler) Problemstrukturen abstrahieren zu können, denn das macht am Ende einen Experten aus.

Je erfahrener man auf einem bestimmten Interessengebiet ist, desto leichter ist es, gute Analogien zu finden und Aufgaben zu lösen – so jedenfalls möchte man meinen. Doch das ist, wie sich herausstellt, nur die halbe Geschichte. Relevant sind nämlich nicht nur strukturelle Aspekte, sondern auch Ähnlichkeits *vergleiche*.

Gick und Holyoak (1980) konnten den Lösungserfolg bei Dunckers Röntgenproblem steigern, indem sie den Probanden mehr als eine Analogie zu lesen gaben (z. B. das Glühbirnenproblem oder die Burggeschichte) und sie aufforderten, konkret zu beschreiben, inwiefern sich diese ähnelten. Unter diesen Versuchsbedingungen schnellte die Erfolgsrate auf über 75 % in die Höhe. Auch Kurtz und Loewenstein (2007) legten ihren Probanden verschiedene Aufgaben vor, die sie vor oder parallel zum Röntgenproblem lösen sollten, und kamen zu demselben Ergebnis: Die Probanden erkannten die strukturellen Ähnlichkeiten zwischen dem Röntgenproblem und den parallel gestellten Aufgaben sehr viel öfter (und kamen beim Röntgenproblem auf die Konvergenzlösung), wenn sie explizit aufgefordert waren, die Aufgaben miteinander zu vergleichen.

Wie stark der Lösungserfolg von Vergleichen der relationalen Strukturen dirigiert wird, macht Cummins (1992) auf sehr anschauliche Weise deutlich. Im Rahmen seiner Studien präsentierte er Übungen zur Algebra in Form von Textaufgaben und forderte seine Probanden auf, sich die Aufgaben genau durchzulesen, sie in lösungsrelevante Kategorien zu ordnen und zu lösen. Das klappte bei einigen auf Anhieb: Sie lasen die Aufgaben durch, ordneten und lösten sie. Eine zweite Gruppe war aufgefordert, ihre Aufmerksamkeit auf wichtige mathematische Beziehungen in-

nerhalb der Aufgaben zu lenken, um textbezogene Aussagen zu verifizieren (z. B. Die Reise dauerte 10 Stunden – Ja oder Nein?). Eine dritte Gruppe war aufgefordert, gleiche mathematische Beziehungen auf andere Textaufgaben abzubilden (z. B. 10 Stunden verhalten sich zu der Zeit bis zur Ankunft bei der vorgegebenen Reise wie 3 Stunden zu der Zeit, bis ein vorgegebenes Fass gefüllt ist – Ja oder Nein?). Wenn es darum ging, die Aufgaben in lösungsrelevante Kategorien zu ordnen, orientierten sich die Teilnehmer der ersten Gruppe (Aufgaben lesen, ordnen und lösen) und die der zweiten Gruppe (Aussagen verifizieren) an Oberflächenähnlichkeiten. Die der dritten Gruppe aber, die relationale Informationen auf andere Aufgaben abbilden sollten, orientierten sich an mathematischen Strukturen. Dieser Gruppe gelang es auch vergleichsweise öfter, eine Lösung für die Aufgaben zu finden.

Die Ergebnisse dieser Studie machen deutlich, dass wir von unwichtigen Details abstrahieren und uns an lösungsrelevante relationale Informationen erinnern können, wenn wir nach Vergleichsfällen suchen, egal, ob diese mathematische oder rechtliche Strukturen haben. Wenn die Wissensinhalte in unserem Gedächtnis auf diese Weise organisiert sind, haben wir eine sehr viel größere Chance, hilfreiche Analogien abzurufen, wenn es darum geht, neue Aufgaben zu lösen.

Analogie als Kern der Kognition

Psychologische Forscher auf dem Gebiet der Kognition sind wie wissbegierige Kinder, die alles auseinandernehmen, um zu sehen, wie die Dinge funktionieren. Zum Beispiel eine

Armbanduhr: In ihrem kindlichen Forscherdrang bauen sie jedes Teil aus, prüfen es einzeln wie auch in Kombination mit anderen Teilen, bis sie schließlich verstehen, welche Funktion jedes einzelne dieser Teile hat, die alle zusammenwirken müssen, damit ein Objekt entsteht, das die Zeit misst. In einem nächsten Schritt sammeln sie alle Einzelteile wieder auf und setzen sie zusammen, um zu sehen, ob sie auch wirklich verstanden haben, wie man eine funktionierende Uhr zusammenbaut. Wenn nicht, gleicht das zusammengebaute Objekt am Ende wohl eher einem abstrakten Kunstwerk als einem funktionierenden Zeitmessgerät.

Da der Begriff Kognition unter anderem den menschlichen Prozess der Informationsverarbeitung beschreibt, gleicht der Aufbau eines kognitiven Systems eher der Konstruktion einer Digitaluhr als einer mechanischen Armbanduhr. Man braucht die richtigen Funktionsmodule sowie die richtige Programmierung, damit sie funktioniert. Frühe Modelle des analogen Denkens beschreiben Produktionssysteme, die aus Regeln, separaten Registern und zielrelevanten Kausalstrukturen bestehen, die im mentalen Lexikon, insbesondere im Langzeit- und Arbeitsgedächtnis, gespeichert sind. Die verschiedenen Komponenten, die für Analogiebildungen unabdingbar sind, waren als Regeln codiert. Eingehende Informationen (*input*) mussten mit diesen Regeln systematisch abgeglichen und widerstreitende Regeln richtig interpretiert werden, um den Regelkonflikt zu lösen. Damit waren die elementaren Schritte des analogen Denkens und Schlussfolgerns vollzogen. Das erste System, das auf Gentners *structure mapping theory*, dem klassischen Theoriemodell der analogen Strukturübertragung, aufbaute, war ein Computerprogramm namens *structure*

mapping engine, entwickelt von Falkenhainer, Forbus und Gentner (1989). In diesem Produktionssystem waren die von Gentner beschriebenen Konzepte als Regeln codiert. Zu den späteren (Datenbank)Systemen zählen Fergusons MAGI (1994), Kuehnes SEQL (2000) und Bursteins CARL (1986). Jedes dieser Systeme konnte die Reproduktion echter analoger Schlüsse mehr oder weniger erfolgreich umsetzen. Doch die Erfolgsrate schnellte enorm nach oben, als Forscher begannen, kognitive Analogiesysteme zu entwickeln, die neuronalen Netzwerkmodellen nachempfunden waren – künstliche Denker, die der Logik des Menschen folgen, und Informationen so codiert und verarbeiten, wie ein menschliches Gehirn es tut. Eines der einflussreichsten und ehrgeizigsten dieser Systeme ist das Analogiemodell LISA, ein neuronales Netzwerksystem, das von John Hummel erschaffen wurde und der *multi-constraint theory* von Holyoak und Thagard eine konkrete Form gibt (Hummel & Holyoak 2005). Der geniale Gedanke, der diesem System zugrunde liegt, ist der, dass LISA auf jeder Abstraktionsebene nach Ähnlichkeiten sucht, nach Oberflächenmerkmalen ebenso wie nach einfachen Beziehungen und Beziehungen höherer Ordnung, und diese in den Analogieprozess einbezieht. Und das geschieht, indem LISA mit Erkennungsprozessen arbeitet, ähnlich wie das menschliche Gehirn.

Um zu verstehen, warum neuronale Netzwerke die natürliche Kognition so erfolgreich abbilden, muss man so einigermaßen wissen, wie unser Gehirn funktioniert. Die Grundeinheiten zur Informationsverarbeitung sind Nervenzellen, sogenannte Neuronen. Neuronen sind Instrumente der Kommunikation. Sie empfangen, verarbeiten und senden Informationen. Neuronen sind darauf spezi-

alisiert, Informationen aufzunehmen, in elektrochemische Signale umzuwandeln und diese an ihren Bestimmungsort weiterzuleiten. Sobald zum Beispiel Licht auf die Netzhaut des Auges (Retina) trifft, reagieren die Nervenzellen, indem sie ein elektrisches Signal erzeugen, das sie über den Sehnerv an das Gehirn senden. Dieses elektrische Signal wird erzeugt, indem die Durchlässigkeit der Zellwände verändert wird, sodass ein Ionenaustausch stattfinden kann. Jede Nervenzelle bekommt Signale von unzähligen vorgeschalteten Zellen und summiert diese Eingangssignale. Erst wenn ein bestimmter Schwellenwert erreicht wird, sendet die Zelle über einen langen, schlauchartigen Nervenzellfortsatz, das sogenannte Axon, selbst ein Signal – man sagt, sie feuert. Sobald das Signal am Ende des Axons angekommen ist, werden Neurotransmitter als chemische Botenstoffe ausgeschüttet, die eine Lücke überqueren (den synaptischen Spalt) und an Rezeptoren der Nachbarnervenzelle andocken. Unser Gehirn lernt, indem es Synapsen modifiziert, sodass sie mehr oder weniger stark auf Signale von anderen Zellen reagieren.

Im Grunde genommen sind Denkvorgänge elektrochemische Prozesse, die sich über ein Netzwerk aus Nervenzellen ausbreiten, und Lernprozesse (egal welcher Art) modifizieren diese Nervenzellen.

Was wäre, wenn man Maschinen bauen könnte, die auf genau die gleiche Weise funktionieren? Die gibt es schon, und sie heißen neuronale Netze. Neuronale Netze bestehen aus künstlichen Neuronen, die sich wie biologische Neuronen verhalten. Diese künstlichen Neuronen können so programmiert werden, dass sie Begriffe („Glühbirne“ oder „Brücke“), Merkmale („zerbrechlich“ oder „schmal“) und

auch andere Formen des Wissenstransfers („weil") repräsentieren. Künstliche Neuronen sind in Schichten organisiert. Die untere Schicht empfängt Informationen aus der Umgebung (*inputs*), die oberste Schicht sendet Informationen (wie Entscheidungen) an die Umgebung aus (*outputs*), und die mittlere Schicht verarbeitet die *inputs*, um sie in plausible *outputs* zu verwandeln.

Im neuronalen Netz sind die Neuronen miteinander verbunden, wobei die Stärke der Verbindung zwischen zwei Neuronen durch ein *Gewicht* ausgedrückt wird. Ginge es darum, ein neuronales Netz zu schaffen, um die oben beschriebenen Aufgaben zu repräsentieren, so hätten die Merkmale, die das „Glühbirnenproblem" beschreiben, allesamt ein großes Gewicht auf die Neuronen, mit denen sie verbunden sind. Das Gleiche gilt für das Soldatenproblem. Die Verbindungen *zwischen* den Geschichten hingegen hätten sehr geringe Gewichte. Dies bedeutet, dass das neuronale Netz jede Geschichte für sich erkennt, aber nicht auf die Idee kommt, dass sie viel miteinander zu tun haben.

Nun nehmen wir einmal an, dieses künstliche neuronale Netz soll Dunckers Röntgenproblem lösen. Ein Röntgenstrahl ist eine Form der Lichtenergie, insofern würde das Signal „Glühbirne" aktiviert, was bedeutet, dass das Netz an das Glühbirnenproblem erinnert wird. Ein Röntgenstrahl ähnelt außerdem einem Laserstrahl, weshalb auch das Signal „Laser" aktiviert würde, und die Aktivierung beider Signale würde sich quer durch das neuronale Netzwerk ausbreiten. In der Folge würden alle stark verbundenen Neuronen auch stark aktiviert, alle schwach verbundenen hingegen nur schwach. Dies bedeutet, das Netzwerk würde an die Glühbirnengeschichte „erinnert". Die Soldatengeschichte

würde zwar ebenfalls aktiviert, doch wäre der Aktivierungsgrad derart schwach, dass er keinen großen Einfluss auf die nachgeschalteten Verarbeitungsprozesse hätte. In Bezug auf ein menschliches neuronales Netz, würden wir sagen, dass die Soldatengeschichte zwar am Tor zum Bewusstsein anklopft, eintreten aber tut die Glühbirnengeschichte.

Was bedeutet das? Es bedeutet, dass das künstliche neuronale Netz die Analogie zwischen Röntgenproblem und Glühbirnengeschichte erkannt hat. Die Konvergenzlösung aus der Glühbirnengeschichte wird aktiviert und steht damit auch für die Lösung des Röntgenproblems zur Verfügung. Und so haben wir durch die Verschaltung von ein paar wenigen einfachen Einheiten (künstliche Neuronen) und Prozessen Gedächtnis, Bewusstsein und analoges Denken modelliert.

Wäre dem Netz lediglich die Soldatengeschichte bekannt gewesen, hätte es wohl kaum Verbindungen zwischen den beiden Geschichten gefunden. Es wäre allenfalls zu einer schwachen Aktivierung der Soldatengeschichte gekommen, die jedenfalls nicht ausgereicht hätte, die abschließende Entscheidung des neuronalen Netzwerks zu beeinflussen. Und so wäre das künstliche neuronale Netz, genau wie das menschliche neuronale Netz, nicht an die Soldatengeschichte erinnert worden, selbst wenn im Gedächtnis eine noch so gute Analogie zur Lösung des Röntgenproblems abrufbereit angelegt war.

Genau wie menschliche Gehirne lernen künstliche neuronale Netze, indem sie die Gewichte der neuronalen Verbindungen modifizieren – und bringen die Neuronen mehr oder weniger zum „Feuern“, wenn sie durch Signale anderer Neuronen aktiviert werden. Angenommen, wir wollen dem Netzwerk sagen: „He, beim Glühbirnenproblem und beim

Soldatenproblem handelt es sich um dieselbe Art von Problem, denn auf beide lässt sich kann die Konvergenzlösung anwenden." Diese Information würde das neuronale Netz zwingen, die beiden Probleme zu vergleichen. Es wird die Verbindungen „belohnen", die an der „Konvergenz" beteiligt waren, indem es das Gewicht zwischen den jeweiligen neuronalen Verbindungen verstärkt, und die Verbindungen „bestrafen", die versagt haben, die Ähnlichkeit zu erkennen. Und das sind diejenigen, die für die neuronale Repräsentation der Oberflächenmerkmale relevant sind. Auf dieser Ebene haben die beiden Probleme nichts gemeinsam, weshalb das Gewicht zwischen diesen Verbindungen deutlich verringert wird. Das Netzwerk hat also die beiden Probleme in seinem (Arbeits) Gedächtnis miteinander verbunden, indem es von nicht relevanten Details abstrahiert hat. Wenn es nun mit dem Röntgenproblem konfrontiert wird, werden die verbliebenen Relationen höherer Ordnung auf die übrigen Repräsentationen beider Analogien abgebildet, und das Netzwerk „erkennt" die Ähnlichkeit zwischen allen drei Problemen.

Man beachte, dass wir keine speziellen Regeln oder Strategien programmieren mussten, um diesem neuronalen System Analogieschlüsse abzugewinnen. Neuronale Netze funktionieren nun einmal so (sie können komplexe Muster wie „Ähnlichkeitsmuster" lernen, ohne vorher die ihnen zugrunde liegenden Regeln entwickelt zu haben). Sie suchen nach Ähnlichkeitsmerkmalen, und Ähnlichkeiten sind auf vielen verschiedenen Ebenen der abstrakten Informationsverarbeitung zu finden. Genau wie das Gehirn, das biologische neuronale Netz des Menschen, sind künstliche neuronale Netze informationsverarbeitende Systeme, die sich Informationen in Form der Aktivierung der Zellen über

ihre Verbindungen zusenden. Genau wie im menschlichen Gehirn breitet sich diese Aktivierung über das Netzwerk aus. Genau wie im Gehirn modellieren die stark aktivierten Verbindungen das Arbeitsgedächtnis. Und genau wie das Gehirn, sind künstliche neuronale Netze lernfähige Systeme, die lernen, indem sie neuronale Verbindungsgewichte verändern. Die natürlich und automatisch ablaufenden Vorgänge führen in beiden Systemen zu Analogieschlüssen. (Einspruch: Natürlich haben künstliche Systeme kein Ich-Bewusstsein; sie sind sich ihrer selbst nicht bewusst. Aber wenn sie das wären, könnte man das, was Ihnen gerade „durch den Kopf" geht, einfach auf technischem Wege bildlich sichtbar machen, nur indem man feststellt, welche Neuronen gerade stark aktiviert sind.)

Es gibt eine weitere äußerst nützliche Eigenschaft, die dem menschlichen Gehirn und künstlichen neuronalen Netzen gemeinsam ist: Beide lernen aus Beispielen. Und beide lernen auf diese Weise am besten. Das bedeutet, dass Beispiele für den Lernerfolg von zentraler Bedeutung sind. Wie jeder gute Lehrer weiß, kann ein einziges gutes Beispiel eine ganze Unterrichtsstunde ersetzen. Und das erleichtert die Arbeit eines Programmierers oder Informatikers ungemein. Er muss nicht erst Objekt- oder Ereigniskategorien implementieren. Er muss das künstliche neuronale System einfach nur mit einer Menge Beispiele „füttern", die das System auf der Basis von Ähnlichkeiten dann selbst zu klassifizieren versucht, um ihm anschließend zu sagen, ob die vorgenommene Klassifizierung richtig oder falsch ist. Den Rest erledigt das System von alleine. Neuronale Netzwerke können sogar ihre eigenen Algorithmen erzeugen, um künftige Aufgaben besser ausführen zu können.

Wie aber steht es mit der „impliziten“ Aktivierung, die nicht stark genug ist, um menschliche Denkprozesse auf einer Bewusstseinsebene nachzubilden? Im vorangegangenen Kapitel haben wir gesehen, dass durch Einsicht gewonnene Lösungen oft als eine schwache Aktivierung in nicht dominanten Hirnarealen beginnen. Neuronale Netzwerke können ohne Weiteres solche Lösungsmethoden entwickeln, die „plötzliche Geistesblitze“ erklären, welche scheinbar wie aus dem Nichts ins Bewusstsein kommen und mitunter ganze Wissensdisziplinen aus den Fugen bringen. Hier einige Beispiele aus der Wissenschaftsgeschichte:

- Friedrich August Kekule von Stradonitz (1829–1896), deutscher Chemiker und Naturwissenschaftler, suchte verzweifelt nach der chemischen Zusammensetzung von Benzol (auch Benzen genannt), bis ihm ein Traum die langgesuchte Lösung brachte: Eine Schlange erschien, biss sich selbst in den Schwanz. Die Erkenntnis traf ihn wie ein Blitz: So wie die sich im Kreise drehende Schlange reihen sich die einzelnen Kohlenstoffatome des Benzolmoleküls in einer Ringstruktur ohne Anfang und Ende aneinander.
- Dmitri Mendelejew (1834–1907), russischer Chemiker, träumte von einem „Tisch, auf dem sich alle Elemente wie erforderlich zusammenfügten“[1] – eine Einsicht, die ihn auf das Periodensystem brachte, in dem alle chemischen Elemente nach Gewicht und Valenz tabellarisch angeordnet sind.

[1] Schwedt G (2009) Noch mehr Experimente mit Supermarktprodukten – Das Periodensystem als Wegweiser. Wiley, Weinheim

- Otto Loewi (1873–1961) glaubte an eine chemische Übertragung der Nervenimpulse, doch es gelang ihm nicht, diese These experimentell nachzuweisen, bis ihm das eigentlich simple Experiment im Traum erschien. Loewi schrieb es aufgeregt auf einen kleinen Zettel auf seinem Nachttisch und schlief wieder ein. Am nächsten Morgen stellte er zu seinem Entsetzen fest, dass er gar nicht lesen konnte, was er da gekritzelt hatte, und sich auch nicht mehr an seinen Traum erinnerte! Zum Glück hatte er den gleichen Traum in der folgenden Nacht wieder. Jetzt überließ er nichts mehr dem Zufall, begab sich direkt in sein Labor und führte das Experiment erfolgreich durch, für das er mit dem Nobelpreis für Medizin ausgezeichnet wurde.

Wir alle erleben tagtäglich analogiebasierte Erkenntnisse, denn „analoges Denken ist die Grundlage für normale Gedächtnisleistungen des Menschen", wie Holyoak und Thagard (1997) formulieren.

Box 9.1 Einige der besten (und witzigsten) Analogiestilblüten von Schülern

Die *Washington Post* schrieb einen Wettbewerb aus, für den Lehrer die komischsten Stilblüten einsenden konnten, auf die sie über die Jahre hinweg in Aufsätzen gestoßen waren. Hier einige der originellsten:

1. Er verfiel ihr mit Haut und Haar, so, als hätte der East River sein Herz verschluckt.
2. Der Plan war ganz einfach, so wie mein Schwager Phil. Nur, dass er im Gegensatz zu meinem Schwager Phil tatsächlich funktionieren könnte.

3. Es kam die Treppe herunter und sah so ähnlich aus wie etwas, das nie ein Mensch zuvor gesehen hatte.
4. Der Löwenzahn wiegte sich im sanften Wind wie ein oszillierender Ventilator auf mittlerer Stufe.
5. Ihre Lippen waren rot und voll wie Blutprobenröhrchen, die ein unkonzentrierter Phlebologe bis zum Anschlag gefüllt hatte.
6. Ihr Vokabular ließ zu wünschen übrig wie noch was.
7. Sie blühte an seiner Seite auf wie eine Kolonie Colibakterien auf einem abgestandenen Stück kanadischen Rindfleischs.
8. Die Lampe stand einfach da wie ein unbelebtes Objekt.
9. Es war eine amerikanische Tradition, wie Väter, die ihre Kinder mit Elektrowerkzeugen durch die Gegend jagen.
10. Sie war todunglücklich, so, wie wenn jemand deine Torte hinaus in den strömenden Regen stellt, die ganze grüne Zuckerglasur verläuft, du noch dazu das Tortenrezept verlierst und sowieso nicht die Bohne singen kannst.
11. Die Gedanken wirbelten nur so in seinem Kopf wie Unterhosen in einem Wäschetrockner, klebten aneinander und fielen wieder auseinander.
12. Er war so groß wie ein fast zwei Meter hoher Baum.
13. Ihr Gesicht war ein vollkommenes Oval, wie ein Kreis, der an beiden Seiten durch ein Oberschenkeltrainingsgerät sanft zusammengedrückt wurde.
14. Vom Dachboden ertönte ein unheimliches Heulen. Alles war unheimlich und unwirklich, so, wie wenn man Urlaub in einer anderen Stadt macht und im Fernsehen *Jeopardy* plötzlich um sieben, statt wie gewohnt um halb acht läuft.
15. John und Mary waren sich noch nie zuvor begegnet, so, wie zwei Kolibris, die sich ebenfalls noch nie begegnet waren.
16. Sie hatte ein tiefes, kehliges, unverfälschtes Lachen, so wie ein Hund klingt, bevor er kotzt.

17. Er war lahm wie eine Ente. Aber nicht wie die sprichwörtliche lahme Ente, sondern wie eine richtige Ente, die tatsächlich lahmt, weil sie vielleicht auf eine Landmine oder so etwas getreten war.
18. Lange Zeit auseinandergerissen durch ein grausames Schicksal, stürmten die beiden Liebenden nun über das grasige Feld aufeinander zu, wie zwei Güterzüge, der eine mit 80 Kilometern pro Stunde aus Cleveland kommend, wo er um 18:36 Uhr abgefahren war, der andere mit 50 Kilometern pro Stunde aus Topeka kommend, wo er um 16:19 Uhr abgefahren war.
19. Der junge Krieger hatte einen hungrig gierigen Blick, so einen, den man kriegt, wenn man lange Zeit nichts gegessen hat.
20. Die Sardinen waren so eng zusammengequetscht wie der Mittelsitz einer Boeing 747.

Literatur

Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *54*, 299–352.

Anderson, J. L. (1998). Embracing uncertainty: The interface of Bayesian statistics and cognitive psychology. *Conservation Ecology*, *2*(1): 2. Available online at http://www.consecol.org/vol2/iss1/art2/

Armstrong, D. (1983). *What is a law of nature*? Cambridge: Cambridge University Press.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.

Bacon, F. (1620). *Novum Organum*. Available online at http://www.constitution.org/bacon/nov_org.htm

Ball, W. (1973). The perception of causality in the infant. Presented at the Meeting of the Society for Research in Child Development, Philadelphia, PA.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chance. *Philosophical Transactions of the Royal Society of London*, *53*(10), 370–418.

Bentham, J. (1789/2005). *Utilitarianism in principles of morals and legislation*. Adamant Media Corporation.

Borg, J. S., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5,

in moral judgments: An fMRI investigation. *Journal of Cognitive Neuroscience*, *18*, 803–817.

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Science*, *9*, 322–328.

Bowers, K. S., Rehehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, *22*, 72–110.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*, 619–639.

Buist, D. S., Anderson, M. L., Haneuse, S. J., Sickles, E. A., Smith, R. A., et al. (2011). Influence of annual interpretive volume on screening mammography performance in the United States. *Radiology*, *259*, 72–84.

Burstein, M. (1986). Concept formation by incremental analogical reasoning and debugging. In R. Michalski et al. (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2, pp. 351–370). New York: Morgan Kaufman.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: a hierarchical approach. *Behavioral & Brain Sciences*, *21*, 667–721.

Camerer, C. F. (2003). Behavioral studies of strategic thinking in games. *Trends in Cognitive Sciences*, *7*, 225–231.

Cartwright, N. (1980). Do the laws of physics state the facts? *Pacific Philosophical Quarterly*, *61*, 75–84.

Chan, D., & Chua, F. (1994). Suppression of valid inferences: Syntactic views, mental models, and relative salience. *Cognition*, *53*, 217–238.

Chase, V. M., Hertwig, R., & Gigerenzer, G. (1998). Visions of rationality. *Trends in Cognitive Sciences*, *2*, 206–214.

Chase, W., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Cheney, D. L., & Seyfarth, R. M. (1992). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365–382.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.

Christiansen, C. L., Wang, F., Barton, M. B., Kreuter, W., Elmore, J. G., Gelfand, A. E., & Fletcher, S. W. (2000). Predicting the cumulative risk of false-positive mammograms. *Journal of the National Cancer Institute*, *92*, 1657–1666.

Cummins, D. D. (1992). Role of analogical reasoning in the induction of problem categories. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *18*, 1103–1124.

Cummins, D. D. (1995). Naive theories and causal deduction. *Memory & Cognition*, *23*, 646–658.

Cummins, D. D. (1997). Reply to Fairley and Manktelow's comment on "Naïve theories and causal deduction." *Memory & Cognition*, *25*, 415–416.

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, *19*, 274–282.

Dawkins, S. (1976). *The selfish gene*. Oxford: Oxford University Press.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*, 684–687.

de Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). Causal conditional reasoning and semantic memory retrieval: A test of

the "semantic memory framework." *Memory & Cognition*, *30*, 908–920.

de Neys, W., Schaeken, W., & d'Ydewalle, G. (2003). Inference suppression and semantic memory retrieval: Every counterexample counts. *Memory & Cognition*, *31*, 581–595.

de Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science*, *19*, 483–489.

de Quervain, D. J, Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, *305*, 1254–1258.

de Waal, F. (1989). Food sharing and reciprocal obligations among chimpanzees. *Journal of Human Evolution*, *18*, 433–459.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, *58*, Whole No. 270.

Eckel, C. C., & Grossman, P. J. (1995). Altruism in anonymous dictator games. *Games and Economic Behavior*, *16*, 181–191.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgements under uncertainty: Heuristics and biases* (pp. 249–267). Cambridge: Cambridge University Press.

Edwards, W. (1955). Experimental measurement of utility. *Econometrica*, *23*, 346–347.

Elio, R. (1998). How to disbelieve p –> q: Resolving contradictions. *Proceedings of the Twentieth Meeting of the Cognitive Science Society*, 315–320.

Evans, J. St. B., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*, 295–306.

Evans, J. St. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual process theory of reasoning. *Thinking & Reasoning*, *11*, 382–389.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine. *Artificial Intelligence*, *41*, 1–63.

Fehr, E., & Fischbacher, U. (2004a). Social norms and human cooperation. *TRENDS in Cognitive Sciences*, *8*, 185–190.

Fehr, E., & Fischbacher, U. (2004b).Third-party punishment and social norms. *Evolution & Human Behavior*, *25*, 63–87.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*, 980–994.

Fehr, E., & Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, *422*, 137–140.

Ferguson, R. W. (1994). MAGI: Analogy-based encoding using symmetry and regularity. In *Proceedings of the 16th Annual Conference of Cognitive Science Society* (pp. 283–288). Mahwah, NJ: Lawrence Erlbaum Associates.

Fiddick, L., & Cummins, D. D. (2007). Are perceptions of fairness relationship specific? The case of noblesse oblige. *Quarter Journal of Experimental Psychology*, *60*, 6–31.

Financial Crisis Inquiry Commission (2011). *The Financial Crisis Inquiry Report, Authorized Edition: Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*. Jackson, TN: Public Affairs.

Foot, P. (1978). *The problem of abortion and the doctrine of the double effect in virtues and vices*. Oxford: Basil Blackwell.

Fugelsang, J., & Dunbar, K. (2005). Brain-based mechanisms underlying complex causal thinking. *Neuropsychologia*, *43*, 1204–1213.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

Gick, M. L., and Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306–355.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is relevant for psychology (and vice

versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129–161). New York: Wiley.

Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you*. New York: Simon & Schuster.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2008). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, *8*, 53–96.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 513–525.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704.

Gigerenzer, G., Todd, P., & the ABC group (2000). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Gobet, F., & Simon, H. A. (1996). The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grand-master-level chess. *Psychological Science*, *7*, 52–55.

Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition*, *87*, B11–B22.

Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences*, *11*, 322–323.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*, 389–400.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–2108.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Strength and structure in causal induction. *Cognitive Psychology*, *51*, 334–384.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*, 98–116.

Haidt, J., & Joseph, C. (2008). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind Volume 3: Foundations and the future* (pp. 367–391). New York: Oxford University Press.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation in preverbal infants. *Nature*, *450*, 557–559.

Harris, L. (2007). *The suicide of reason: Radical Islam's threat to the West*. New York: Basic Books.

Heekeren, H. R., Wartenburger, I., Schmidt, H., Schwintowski, H., & Villringer, A. (2003). An fMRI study of simple ethical decision-making. *Neuroreport*, *14*, 1215–1219.

Hertwig, R., & Gigerenzer, G. (1999). The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, *12*, 275–305.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games & Economic Behavior*, *7*, 346–380.

Hoffman, E., McCabe, K., & Smith, V. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review*, *86*, 653–660.

Hoffman, E., & Spitzer, M. (1985). Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *Journal of Legal Studies*, *15*, 254–297.

Hofman, W., & Baumert, A. (2010). Immediate affect as a basis for intuitive moral judgement: An adaptation of the affect misattribution procedure. *Cognition & Emotion*, *24*, 522–535.

Hofstadter, D. (2009). Analogy as the core of cognition. In D. Gentner, K. Holyoak & B. Kokinov (Eds.), *The Analogical mind: Perspectives from cognitive science* (pp. 499–538). Cambridge, MA: The MIT Press.

Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the eventrelated brain potential associated with reward and punishment. *Psychophysiology*, *41*, 245–253.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, *15*, 332–340.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*, 295–355.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist*, *52*, 35–44.

Hume, D. (1740/1967). *A treatise of human nature*. Oxford: Oxford University Press.

Hume, D. (1748/2010). *An enquiry concerning human understanding*. New York: General Books.

Hume, D. (1751/1998). *An enquiry concerning the principles of morals*. Oxford: Oxford University Press.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*, 427–466.

Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture: An overview of the LISA project. *Current Directions in Cognitive Science*, *14*, 153–157.

Inglis, M., & Simpson, A. (2007). Belief bias and the study of mathematics. In D. Pitta-Pantazi & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 2310–2319). Larnaca, Cyprus.

Janveau-Brennan, G., & Markovits, H. (1999). The development of reasoning with causal conditionals. *Developmental Psychology*, *35*, 904–911.

Juslin, P., Winman, A., & Olsson, H. (2000). Naïve empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review*, *107*, 384–396.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kant, I. (1781/2008). *Critique of pure reason*. New York: Penguin Classics.

Kant, I. (1783/1911). *Prolegomena to any future metaphysics*. Akademie edition, vol. IV, Berlin.

Kant, I. (1785/1989). *The foundations of the metaphysics of morals*. Upper Saddle River, NJ: Prentice-Hall.

Kant, I. (1787/1997). *The critique of practical reason*. Cambridge: Cambridge University Press.

Kaufman, J., & Zigler, E. F. (1988). Do abused children become abusive parents? *Annual Progress in Child Psychiatry & Child Development*, *29*, 591–600.

Keller, G. (2011). *Statistics for management and economics*. Chula Vista, CA: Southwestern College Publishing.

Kelley, H. H., & Stahelski, A. J. (1970). The inference of intention from moves in the Prisoner's Dilemma Game. *Journal of experimental social psychology*, *6*, 401–419.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, *308*, 78–83.

Kirsch, I., Deacon, B. J., Huedo-Medina, T. B., Scoboria, A., Moore, T. J., & Johnson, B. T. (2008). Initial severity and antidepressant benefits: A meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine*, 5(2).

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Knutson, B., & Bossaerts, P. (2007). Neural antecedents of financial decisions. *Journal of Neuroscience*, *27*, 8174–8177.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *Journal of Neuroscience*, *25*, 4806–4812.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, *446*, 908–911.

Kosfeld, M., Heinrichs, M., Zaks, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*, 673–676.

Kotovsky, L., & Baillargeon, R. (2000). Reasoning about collision events involving inert objects in 7.5-month-old infants. *Developmental Science*, *3*, 344–359.

Kuehne, S. E., Forbus, K. D., Gentner, D., & Quinn, B. (2000). SEQL: Category learning as progressive abstraction using structure mapping. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Kurtz, K. J., & Loewenstein, J. (2007). Converging on a new role for analogy in problem solving and retrieval: When two problems are better than one. *Memory & Cognition*, *35*, 334–341.

Lakshminarayanan, V. R., Chen, M. K., & Santos, L. (2011). The evolution of decision-making under risk: Framing effects in monkey risk preferences. *Journal of Experimental Social Psychology*, *47*, 689–693.

Lewis, D. (1973). Causation. Journal of Philosophy, 70, 556–567.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Nous*, *13*, 455–476.

Lewis, D. (2000). Causation as influence (abridged version). *Journal of Philosophy*, *97*, 182–197.

Lewis, D. (2004). Void and object. In J. Collins, N. Hall & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 277–290). Cambridge, MA: MIT Press.

Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–344). Cambridge/New York: Cambridge University Press.

Lombrozo, T. (2009). The role of moral commitments in moral judgment. *Cognitive Science*, *33*, 273–286.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

Luchins, A. (1942). *Mechanization in problem solving. Psychological Monographs 34*. Washington, DC: American Psychological Association.

Luchins, A., & Luchins, E. H. (1959). Rigidity of behavior – a variational approach to the effect of einstellung. Eugene: University of Oregon Books.

Luo, J., Yuan, J., Qiu, J., Zhang, Q., Zhong, J., & Huai, Z. (2008). Neural correlates of the belief-bias effect in syllogistic reasoning: An event-related potential study. *NeuroReport*, *19*, 1075–1080.

Mackie, J. L. (1974). *The cement of the universe: A study in causation*. Oxford: Clarendon.

Martin, J. A., & Elmer, E. (1992). Battered children grown up: A follow-up study of individuals severely maltreated as children. *Child Abuse & Neglect*, *16*, 75–88.

Maynard Smith, J. (1972). *On evolution*. Edinburgh: Edinburgh University Press.

McCarthy, J. (1980). Circumscription – a form of non-monotonic reasoning. *Artificial Intelligence*, *13*, 27–39.

Metcalfe, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 288–294.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and non-insight problem solving. *Memory & Cognition*, *15*, 238–246.

Michotte, A. (1963). *The perception of causality*. New York: Basic Books.

Mill, J. S. (1843/1963). *System of logic, ratiocinative and inductive*. Reprinted in J. M. Robson (Ed.), *Collected Works of John Stuart Mill*. Toronto: University of Toronto Press.

Mill, J. S. (1859/2011). *On liberty*. Hollywood, FL: Simon Brown.

Mill, J. S. (1861/2007). *Utilitarianism*. Mineola, NY: Dover.

Mill, J. S. (1869/2011). *Subjection of women*. Clippesby: Croft Classics.

Moll, E., & de Oliveira-Souza (2001). Frontopolar and anterior temporal cortex activation in a moral judgment task: Preliminary function MRI results in normal subjects. *Arq Neuropsiquiatr*, *59*, 657–664.

Morgenstern, O., & von Neumann, J. (1944). Theory of games and economic behavior. Princeton, NJ: Princeton University Press.

Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive Psychology*, *61*, 63–86.

Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36(1)*, 48–49.

National Cancer Institute Surveillance, Epidemiology, and End Results (SEER). Retrieved from *http://seer.cancer.gov/statfacts/html/breast.html*

Newborn, M. (1997). Kasparov versus Deep Blue: Computer chess comes of age. New York: Springer.

Newell, A., & Simon, H. (1963). GPS: A program that simulates human thought. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought* (pp. 279–293). New York: McGraw-Hill.

Oaksford, M., & Chater, N. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Science*, *3*, 57–65.

Oaksford, M., Chater, N., Grainger, B., & Larkin, J. (1997). Optimal data selection in the reduced array selection task (RAST). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 441–458.

O'Doherty J., Deichmann R., Critchley H. D., & Dolan R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, *33*, 815–826.

Olsen, V. (2004). Man of action. *Smithsonian Magazine*, September. Retrieved from http://www.smithsonian.com

O'Rourke, P. J. (1999). *Eat the rich*. London: Atlantic Monthly Press.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*, 96–146.

Peirce, C. (1877). The fixation of belief. *Popular Science Monthly*, 1–15.

Pollock, J. (1987). Defeasible reasoning. *Cognitive Science*, *11*, 481–518.

Polya, G. (1945). *How to solve it*. Princeton, NJ: Princeton University Press.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American. Economics Review*, *83*, 1281–1302.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*, 48–55.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron*, *35*, 395–405.

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology*, *16*, 371–416.

Russell, B. (1959). *My philosophical development*. London/New York: George Allen and Unwin/Simon and Schuster.

Sanfey, A. G. (2007). Social decision-making: Insights from game theory and neuroscience. *Science*, *318*, 598–602.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, *300*, 1755–1758.

Schnall, S., Benton, J., & Harvey, S. (2008a). With a clean conscience: Cleanliness reduces the severity of moral judgments. *Psychological Science*, *19*, 1219–1222.

Schnall, S., Haidt, J., Clore, G., & Jordan, A. (2008b). Disgust as embodied moral judgment. *Personality & Social Psychology Bulletin*, *34*, 1096–1109.

Schwitzgebel, E., & Cushman, F. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language*, *27*, 135–153.

Silveira, J. (1971). Incubation: The effect of interruption timing and length on problem solution and quality of problem processing. Reported in J. R. Anderson (1985). *Cognitive psychology and its implications*. Basingstoke: W. H. Freeman.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior & Human Decision Processes*, *91*, 296–309.

Smith, A. McCall (2007). *Love over Scotland*. New York: Anchor Books.

Sober, E., & Sloan-Wilson, D. (1999). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.

Sohn, E. (2001). Stopping time in its tracks. *U.S. News & World Report*, July 9. (Also electronically reprinted on www.usnews.com)

Spiro, M. E. (1996). Postmodernist anthropology, subjectivity, and science: A modernist critique. In *Comparative Studies in Society and History* (Vol. 5, pp. 759–780). Ann Arbor: University of Michigan Press.

Thomson, J. J. (1985). The trolley problem. *Yale Law Journal*, *94*, 1395–1415.

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211, 453–458.

Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84–98). New York: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315.

U.S. Securities and Exchange Commission and the Commodity Futures Trading Commission (2010). Findings regarding the market events of May 6, 2010. September 30.

van Dijk, E., & Vermunt, R. (2000). Strategy and fairness in social decision making: Sometimes it pays to be powerless. *Journal of Experimental Social Psychology*, *36*, 1–25.

Vershueren, N., Schaeken, W., de Neys, W., & d'Ydewalle, G. (2004). The difference between generating counterexamples and using them during reasoning. *Quarterly Journal of Experimental Psychology*, *57*, 1285–1308.

U.S. Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*, 273–281.

Weg, E., & Smith, V. (1993). On the failure to induce meager offers in ultimatum games. *Journal of Economic Psychology*, *14*, 17–32.

Westin, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 U.S. presidential election. *Journal of Cognitive Neuroscience*, *18*, 1947–1958.

White, P. A. (1995). Use of prior beliefs in the assignment of causal roles: causal powers versus regularity-based accounts. *Memory & Cognition*, *23*, 43–54.

White, P. A. (2000). Causal judgment from contingency information: The interpretation of factors common to all instances. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1083–1102.

Whitehead, A. N., & Russell, B. (1910, 1912, 1913). *Principia mathematica* (3 vols). Cambridge: Cambridge University Press.

Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, *308*, 181–184.

Williams, A. M., Davids, K., Burwitz, L., & Williams, J. G. (1992). Perception and action in sport. *Journal of Human Movement Studies*, *22*, 147–204.

Williams, A. M., Davids, K., Burwitz, L., & Williams, J. G. (1994). Visual search strategies in experienced and inexperienced soccer players. *Research Quarterly for Exercise and Sport*, *65*, 127–135.

Wilson, W. A., & Kuhn, C. M. (2005). How addiction hijacks our reward system. *Cerebrum*, *7*, 53–66.

Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*, 333–362.

Young, L., Camprodon, J. A., Hauser, M., Pasual-Leones, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*, 6753–6758.

Zacks, R., & Hasher, L. (2002). Frequency processing: A twenty-five-year perspective. In P. Sedlmeier & T. Betsch (Eds.), *Frequency processing and cognition* (pp. 21–36). New York: Oxford University Press.

Zak, P. J., Stanton, A. A., & Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLoS ONE 2* (11): e1128. doi:10.1371/journal.pone.0001128.

Sachverzeichnis

D. D. Cummins, *Gutes Denken,* DOI 10.1007/978-3-642-41810-5,