

Plausibilität, Wahrscheinlichkeit, Kohärenz

Mark Siebel

In den letzten Jahren sind wieder verstärkt Bemühungen unternommen worden, erkenntnis- und wissenschaftstheoretische Fragen sowie die zugehörigen Begrifflichkeiten einer wahrscheinlichkeitstheoretischen Behandlung zuzuführen. Colin Howsons und Peter Urbachs Buch *Scientific Reasoning. The Bayesian Approach* (1993) ist einer der Auslöser, an neueren Monografien sind Branden Fitelsons *Studies in Bayesian Confirmation Theory* (2001) sowie Luc Bovens' und Stephan Hartmanns *Bayesian Epistemology* (2003) zu nennen. Daneben gibt es eine Vielzahl von Aufsätzen, die sich aus der Perspektive des Wahrscheinlichkeitskalküls mit Themen wie Theorienbewertung, Bestätigung, Verlässlichkeit, Zeugenaussagen und Kohärenz beschäftigen. Tomoji Shogenji (1999), Erik Olsson (2002) und Fitelson (2003) haben beispielsweise Vorschläge unterbreitet, wie sich die Kohärenz eines Aussagensystems probabilistisch berechnen lässt.¹

Wie schon die Titel der Bücher deutlich machen, bezeichnen sich die Protagonisten dieser Bewegung gern als Bayesianer. Das kann zu Missverständnissen führen, da viele mit dem Stichwort „Bayesianismus“ nicht nur die Anwendung der Wahrscheinlichkeitsrechnung auf epistemologische Angelegenheiten assoziieren, sondern auch die Idee, dass Wahrscheinlichkeiten rationale Überzeugungsgrade repräsentieren. Dem würde aber längst nicht jeder Freund der probabilistischen Erkenntnistheorie zustimmen. Ich werde mich deshalb des weniger verfänglichen Terminus „Probabilismus“ bedienen.

Der Probabilismus ist aus den verschiedensten Gründen kritisiert worden. So sieht er sich der Frage ausgesetzt, ob es in realen Situationen überhaupt möglich ist, die benötigten Wahrscheinlichkeiten in einer vernünftigen Weise zu fixieren. Große Schwierigkeiten bereiten etwa neue wissenschaftliche Hypothesen. Ihnen lassen sich vielleicht subjektive Überzeugungsgrade zuordnen, aber diese im engen Sinn bayesianische Herangehensweise hat einen unschönen Beigeschmack von Beliebigkeit.

Meine Kritik geht in eine andere Richtung. Ich werde zu zeigen versuchen, dass die Art und Weise, in der der Probabilist die Plausibilität von Hypothesen beurteilt, zu eingeschränkt ist: Er sieht nur auf ihre Wahrscheinlichkeiten, während tatsächlich auch andere Faktoren – insbesondere ihr Erklärungspotential – eine gewichtige Rolle spielen (Abschnitte 1 bis 3). Eines der Beispiele, mit denen die probabilistische Erkenntnistheorie Schwierigkeiten hat, stammt aus den Experimenten zu unsicheren Schlüssen, die von den Psychologen Amos Tversky und Daniel Kahneman unternommen wurden. Ich werde in Abschnitt 4 vorführen, dass ein kohärenztheoretischer Ansatz dieses Beispiel in den Griff bekommt. Davon ausgehend wird im Schlussabschnitt 5 ein neuer Vorschlag gemacht, wie sich die Ergebnisse dieser Experimente erklären lassen. Und es wird darauf

¹ In Siebel 2004a wird Fitelsons Vorschlag kritisiert; Siebel 2005 setzt sich auch mit Shogenjis und Olssons Kohärenzmaßen sowie der Quasiordnung von Bovens und Hartmann (2003, Kap. 2) auseinander.

verwiesen, dass sich der Vorwurf der Irrationalität, dem Tversky und Kahneman die Versuchspersonen aussetzen, auf dieser Basis nicht halten lässt.

1. Probabilistische Erkenntnistheorie

Elliott Sober (1994, 227) hat die Grundidee der probabilistischen Erkenntnistheorie auf den Punkt gebracht: „Bayesians think that observations confer probabilities on hypotheses and that the mathematical idea of probability is the right way to measure the epistemic property of plausibility.“ Etwas genauer gehen Probabilisten davon aus, dass sich die Plausibilität einer Hypothese H im Lichte der Daten D (und des weiteren Hintergrundes) mit der bedingten Wahrscheinlichkeit $P(H|D)$, der Nachherwahrscheinlichkeit der Hypothese, identifizieren lässt.² H mag eine wissenschaftliche Theorie sein, während D eine Beschreibung experimenteller Ergebnisse ist; es kann sich aber auch um Schuldvermutungen („Der Angeklagte hat den Mord begangen“) und Zeugenaussagen („Ich habe ihn kurz nach dem Mord in der Nähe des Tatortes gesehen“) oder um Allerweltsbehauptungen handeln.

Zur Berechnung von $P(H|D)$ lässt sich eine Folgerung aus den wahrscheinlichkeitstheoretischen Axiomen einsetzen, das Bayes-Theorem:

$$P(H|D) = P(H) \cdot \frac{P(D|H)}{P(D)}$$

Die Ausgangswahrscheinlichkeit $P(H)$ der Hypothese wird dabei analog zu $P(H|D)$ als die Plausibilität gedeutet, die H vor der Auswertung der Daten hatte. $P(D)$ steht für das Ausmaß, in dem die Daten allein auf der Basis des bisherigen Hintergrundes zu erwarten waren. Und die so genannte Likelihood $P(D|H)$ gibt den Grad wieder, zu dem sie zu erwarten sind, wenn die Hypothese als zutreffend vorausgesetzt wird. Da eine bedingte Wahrscheinlichkeit mit 1 anzusetzen ist, wenn das Bedingte aus der Bedingung logisch folgt,³ ist die Likelihood 1, falls die Beobachtungen von der Hypothese und den Hintergrundannahmen impliziert werden.

Der Vorschlag der Probabilisten lautet dann: Liegt $P(H|D)$ nahe bei 1, dann ist H im Lichte von D überzeugend; tendiert diese Wahrscheinlichkeit in Richtung 0, bieten die Daten dagegen einen Grund, die Hypothese zu verwerfen. Und wenn $P(H|D)$ sich im Umfeld von 0,5 befindet, dann geben die Beobachtungen nicht genug her, um auf ihrer Grundlage eine Beurteilung vorzunehmen (vgl. z.B. Jaynes 1996, 404). Auf einen einfachen Fall angewandt: Vor dem (hypothetischen) Hintergrund, dass 90 % aller Hamburger Labskaus mögen, verleiht die Feststellung, dass Eva Hamburgerin ist, der Annahme, dass sie Labskaus mag, einige Plausibilität, da $P(H|D) = 0,9$.

Eine weitere Frage (über die einem $P(H|D)$ allein nichts sagt) ist, ob die Beobachtungen überhaupt zu irgendeiner *Veränderung* der Plausibilität geführt haben: Ist die Hypothese

² Ich lasse den Hintergrund der Einfachheit halber aus den Formeln heraus.

³ Vorausgesetzt wird dabei, dass die Bedingung keinen Widerspruch enthält. Sollte B widersprüchlich sein, dann ist $P(A|B)$ in der Standard-Wahrscheinlichkeitstheorie nicht definiert.

durch der Daten vertrauenswürdiger geworden? Hat sie einen Glaubwürdigkeitsverlust hinnehmen müssen? Oder ist alles beim Alten geblieben? Es hat sich in wissenschaftstheoretischen Kreisen eingebürgert, im ersten Fall von *Stützung* (oder Bestätigung) zu sprechen: Daten stützen eine Annahme, wenn sie ihre Plausibilität heraufsetzen. Dabei wird unter „Plausibilitätserhöhung“ auch die Verringerung von Unplausibilität subsumiert: Experimentergebnisse sprechen ebenfalls für eine Theorie, wenn sie ihre Unglaubwürdigkeit entschärfen. Stützung in diesem Sinn ist somit nicht notwendig damit verbunden, dass die Hypothese auf der Grundlage der Daten plausibel ist. Umgekehrt kann eine Theorie weiter gut dastehen, auch wenn die Beobachtungen ihre Plausibilität herabsetzen.⁴

Mit der Identifikation von Plausibilität und Wahrscheinlichkeit ist dann auch schon festgelegt, wie Stützung probabilistisch zu fassen wäre, nämlich als Wahrscheinlichkeitserhöhung. Entsprechend lautet das so genannte *Relevanzkriterium* (vgl. u.a. Howson und Urbach 1993, 117):

Die Daten stützen die Hypothese gdw. $P(H|D) > P(H)$; sie sprechen gegen sie gdw. $P(H|D) < P(H)$; und sie verhalten sich neutral gdw. $P(H|D) = P(H)$.

In Verbindung mit Bayes' Theorem kann dieses Kriterium einige Pluspunkte für sich verbuchen. Es ergibt sich etwa, dass Theorien gestärkt werden, wenn sich deduktive Konsequenzen aus ihnen nachweisen lassen.⁵ Weniger bekannt dürfte sein, dass der Probabilismus auch im Falle von Experimenten, die irrelevant für eine Theorie sind, weil sie gar keine Prognose der Ergebnisse erlaubt, die richtige Antwort gibt. Dass eine Theorie keine Aussagen über die Daten macht, wäre so zu übersetzen, dass sie deren Wahrscheinlichkeit weder herauf- noch herabsetzt: $P(D|H) = P(D)$. Dann aber muss nach Bayes' Formel auch die Nachherwahrscheinlichkeit der Hypothese identisch mit ihrer Vorherwahrscheinlichkeit sein, was aufgrund des Relevanzprinzips heißt, dass die Hypothese durch die Beobachtungen weder bestätigt noch entkräftet wird.

Das Relevanzkriterium ist ein qualitatives Prinzip: Es macht keine Aussage darüber, wie *stark* Beobachtungen für eine Vermutung sprechen. Was die quantitative Frage angeht, sieht man sich mit einer Unmenge von Vorschlägen konfrontiert. So sind als Maße für die Bestätigungskraft u.a. die Differenz $P(H|D) - P(H)$ sowie das Verhältnis $P(H|D)/P(H)$ oder ein Logarithmus aus ihm empfohlen worden. Diese Entwürfe werden hier nicht diskutiert.⁶ Meine Kritik setzt nicht an der Peripherie, sondern im Zentrum der probabilistischen Erkenntnistheorie an: der Gleichsetzung von Plausibilität und Wahr-

⁴ Carnap (1962, neues Vorwort) hat den Unterschied zwischen Plausibilität und Stützung unter den Stichworten „confirmation as firmness“ und „confirmation as increase in firmness“ verhandelt.

⁵ Vgl. u.a. Earman 1992, 64. Bayesianer behaupten auch, dass sich mit ihrem Ansatz Hempels berüchtigte Rabenparadoxie auflösen lasse. Dass ihre Vorschläge keine wirkliche Lösung bieten, wird in Siebel 2004b zu zeigen versucht.

⁶ Mehr zu den verschiedenen Maßen und ihren Vor- und Nachteilen findet sich in Eells und Fitelson 2002, Fitelson 1999, Fitelson 2001.

scheinlichkeit sowie dem damit verknüpften Relevanzprinzip. Wenn dieses Prinzip keine vernünftige Analyse des Stützungsbegriffs bietet, dann sind auch die Maße, die sich nach ihm richten, nicht akzeptabel.

2. Zum Relevanzkriterium

Es ist eine (mathematisch gesehen) ganz triviale Folgerung aus dem Relevanzkriterium, die Schwierigkeiten bereitet: Daten, die von der Negation einer Hypothese impliziert werden, sprechen gegen sie, solange $P(D)$ kleiner als 1 und größer als 0 ist. Wenn D aus $\neg H$ folgt, ist $P(\neg H \& D)$ identisch mit $P(\neg H)$. Gilt also $0 < P(D) < 1$, so erhalten wir:

$$P(\neg H|D) = \frac{P(\neg H \& D)}{P(D)} = \frac{P(\neg H)}{P(D)} > P(\neg H)$$

Also: $P(H|D) < P(H)$. Laut dem Relevanzprinzip können Daten eine Annahme somit nicht stützen, wenn es nicht sein kann, dass beide falsch sind – kurz: wenn es sich um subkonträre Aussagen handelt. Dieses Ergebnis mag schon für sich betrachtet merkwürdig erscheinen. Ziehen wir aber keine voreiligen Schlüsse, sondern erweitern wir das Blickfeld ein wenig, indem wir dem Begriff der Stützung denjenigen der Kohärenz an die Seite stellen (siehe Siebel 2004c).

Sieht man sich an, wie Wissenschaftler mit diesen Begrifflichkeiten arbeiten, so ist eine Verbindung zwischen ihnen offensichtlich: Eine Theorie bildet nur dann eine kohärente Einheit mit experimentellen Ergebnissen, wenn diese sie stützen. Das bedeutet umgekehrt: Sprechen die Daten *gegen* die Theorie, dann passen sie *nicht* mit ihr zusammen. Man denke etwa an einen Chemiker, der ein Experiment zu Boyles Gasgesetz durchführt. Bevor er (bei konstanter Temperatur) das Volumen des Gases auf die Hälfte reduziert, zeigt das Barometer einen Druck von 1 Bar an, danach sind es 3 Bar. Da Boyles Gesetz nicht 3 Bar, sondern einen nur doppelt so hohen Druck prognostiziert, wird der Chemiker dieses Datum als einen Beleg gegen das Boyle-Gesetz betrachten (sofern er davon ausgeht, dass das Barometer verlässlich arbeitet ...). Und er wird entsprechend konstatieren, dass dieses Gesetz nicht mit dem Resultat seines Experiments harmonisiert. Wissenschaftlich einschlägige Konzeptionen von Stützung und Kohärenz sollten also das folgende *Brückenprinzip* respektieren:

Wenn D gegen H spricht, dann passt H nicht mit D zusammen.

Dieses Prinzip bietet eine Adäquatheitsbedingung für Stützungstheorien. Das heißt nicht, dass eine vernünftige Analyse es *implizieren* sollte, sondern nur, dass sie ihm *nicht widersprechen* darf: Wenn H mit D kohäriert, darf sie nicht nach sich ziehen, dass D H unterminiert. Genau diese Adäquatheitsbedingung erfüllt das Relevanzkriterium aber nicht – und zwar wegen seiner speziellen Behandlung von subkonträren Propositionen.

Nehmen wir an, aus einer Urne werde eine Kugel gezogen. Die Urne enthält acht schwere, weiße Kugeln, eine leichte, weiße sowie eine schwere, schwarze. Wir haben

hier also eine sehr große Überlappung: Von den insgesamt neun weißen Kugeln sind acht schwer, und dasselbe gilt umgekehrt. Aufgrund dieser starken Überschneidung können wir konstatieren, dass das Aussagenpaar

Die gezogene Kugel ist weiß.

Die gezogene Kugel ist schwer.

kohärent ist. In jedem Fall gibt es keinen Grund für die Annahme, dass sie nicht zusammenpassen.

Auf diese Annahme aber sind die Freunde der wahrscheinlichkeitstheoretischen Herangehensweise verpflichtet. Denn erstens sind diese Aussagen subkonträr: Eine von ihnen muss wahr sein, da jede Kugel in der Urne weiß oder schwer ist. Und zweitens besitzen sie beide eine Wahrscheinlichkeit unter 1, weil eine der Kugeln nicht weiß und eine nicht schwer ist. Aus dem Relevanzkriterium ergibt sich deshalb, dass diese Propositionen *gegeneinander* sprechen – was nach dem Brückenprinzip heißt, dass sie ein *inkohärentes* Paar bilden. Da das Brückenprinzip außer Frage steht, bricht somit ein Stützfeiler der probabilistischen Erkenntnistheorie zusammen: Das Relevanzkriterium ist nicht tragfähig, weil es einen zu der absurden Konsequenz führt, dass subkonträre Aussagen keine Kohärenz für sich verbuchen können.

3. Zur Grundidee der probabilistischen Erkenntnistheorie

Die Überlegung des vorherigen Abschnitts wirft nicht nur ein schlechtes Licht auf das Relevanzprinzip, sondern auch auf die grundlegende Idee der Probabilisten: die Gleichsetzung von Plausibilität und Wahrscheinlichkeit. Denn wenn sie in Ordnung wäre, dann müsste die im Relevanzkriterium behauptete Identifikation von Plausibilitäts- und Wahrscheinlichkeitserhöhung ebenfalls unkritisch sein. – Warum aber haut die Grundidee nicht hin?

Die Diagnose ist recht simpel und offensichtlich. Ich will gar nicht leugnen, dass die Wahrscheinlichkeit einer Annahme eine große Rolle spielt, wenn wir sie auf ihre Glaubwürdigkeit hin untersuchen. So kann ich Folgendes ohne Bedenken unterschreiben: Je wahrscheinlicher es ist, dass die Hypothese zutrifft, desto mehr Plausibilität kann sie, *ceteris paribus*, für sich verbuchen. Aber das gilt eben nur mit der Einschränkung „*ceteris paribus*“. Wenn die sonstigen Umstände nicht dieselben sind, dann kann es durchaus vorkommen, dass eine Annahme besser als eine andere dasteht, obwohl sie die kleinere Wahrscheinlichkeit besitzt. Der Grund dafür ist, dass bei der Bewertung von Hypothesen weitere Faktoren neben ihrer Wahrscheinlichkeit einfließen. Einer von ihnen ist jedem Natur- wie Geisteswissenschaftler wohlbekannt und wird auch von Philosophen nicht selten eingesetzt, wenn es darum geht, eine Vermutung als plausibel hinzustellen: ihr *Erklärungspotential*.

Denken wir uns zwei physikalische Hypothesen H_1 und H_2 , die um einen Phänomenbereich konkurrieren, und eine Sammlung D von Daten aus diesem Bereich. $P(H_1|D)$ sei

kleiner als $P(H_2|D)$, dafür aber erklärt H_1 eine größere Teilmenge aus D und bietet darüber hinaus auch schlüssigere Erklärungen als H_2 . Welche Hypothese sollte ein Physiker auswählen? Der gesunde Menschenverstand sagt einem ebenso wie der wissenschaftliche Verstand, dass in so einem Fall eine Abwägung nötig ist. Auch wenn man mit H_1 aufgrund der geringeren Wahrscheinlichkeit ein höheres Risiko eingehen mag, kann dies doch durch das größere Erklärungspotential aufgewogen werden. Schließlich will der Physiker *verstehen*, was in dem Phänomenbereich vor sich geht; und je mehr der Daten eine Theorie erklären kann und je besser die Erklärungen sind, umso tiefer geht sein Verständnis. Wer hier – wie der erkenntnistheoretische Probabilist – nur auf die Wahrscheinlichkeiten der Hypothesen achtet, übersieht, dass es zusätzliche Kriterien gibt, an denen sich ihre Glaubwürdigkeit bemisst.

Dass Probabilisten die Bewertung von Hypothesen zu einseitig ausrichten, wird vielleicht besonders deutlich, wenn man sich Situationen ansieht, in denen die fraglichen Hypothesen nicht miteinander konkurrieren, sondern in denen die eine eine logische Konsequenz der anderen ist (vgl. Glymour 1998, 598). Die probabilistische Beurteilung derartiger Fälle ist durch die wahrscheinlichkeitstheoretischen Axiome eindeutig festgeschrieben. Wie Howson und Urbach (1993, 25) so schön schreiben, respektieren Wahrscheinlichkeiten nämlich die Folgerungsbeziehung: Wenn H_2 aus H_1 folgt, dann ist $P(H_1|D)$ höchstens genauso groß wie $P(H_2|D)$ – und zwar unabhängig davon, wie die Daten D im Detail beschaffen sind. Dagegen ist soweit nichts einzuwenden. Mit der Gleichsetzung von Plausibilität und Wahrscheinlichkeit überträgt der Probabilist diese Eigenschaft dann aber auf Plausibilität. Er muss behaupten: Wie auch immer die Daten aussehen, eine Hypothese kann in ihrem Licht niemals glaubwürdiger sein als Folgerungen aus ihr. Und das ist viel zu unflexibel.⁷

Betrachten wir ein Beispiel, das Tversky und Kahneman in ihren Untersuchungen zu unsicheren Schlüssen eingesetzt haben:⁸ Linda ist 31 Jahre alt, allein stehend, sehr klug und sagt frei heraus ihre Meinung. Sie hat einen Magister in Philosophie gemacht, sich als Studentin intensiv in Fragen der Diskriminierung und der sozialen Gerechtigkeit engagiert und an Anti-Atomkraft-Demonstrationen teilgenommen. Wie stehen die folgenden Behauptungen im Lichte dieser Informationen da? Welche von ihnen kann am ehesten akzeptiert werden? Und wie weit leuchten die anderen ein?

- (B) Linda ist eine Bankangestellte.
- (BF) Linda ist eine Bankangestellte und aktiv in der feministischen Bewegung.
- (F) Linda ist aktiv in der feministischen Bewegung.

Während F ausgesprochen plausibel ist, da aktive Feministinnen durchaus Lindas Merkmale aufweisen können, überzeugt B nicht, weil es nur schwer vorstellbar ist, dass eine

⁷ Wenn es nicht um Plausibilität, sondern um Stützung geht, ist die probabilistische Konzeption flexibler: Die wahrscheinlichkeitstheoretischen Regeln verbieten es nicht, dass $P(H_1|D) > P(H_1)$, während $P(H_2|D) < P(H_2)$, wenn H_1 H_2 impliziert. Mackie (1969, 36) nennt ein Beispiel.

⁸ Vgl. Tversky and Kahneman 1982, 91–94; Kahneman und Tversky 1982, 496.

Bankangestellte den beschriebenen Hintergrund hat. Was aber ist mit BF, der Konjunktion aus B und F? Probabilistisch betrachtet ist sie höchstens genauso plausibel wie B, da eine Konjunktion jedes ihrer Konjunkte impliziert. Es ist deshalb auch vollkommen egal, wie die Informationen aussehen, die wir über Linda bekommen. Ein Probabilist muss in jedem Fall davon ausgehen, dass BF höchstens genauso viel Glaubwürdigkeit für sich verbuchen kann wie B, da $P(BF|D)$ nicht über $P(B|D)$ hinausgehen kann.

Damit verliert der Probabilist aber aus dem Blick, dass BF in einer anderen, nicht minder essentiellen Hinsicht besser als B dasteht: Die Konjunktion enthält mit der Feminismus-Komponente einen Bestandteil, der sich in vielfältige Erklärungsbeziehungen zu Teilen der Daten wie auch zu dem anderen Konjunkt setzen lässt. So kann die Tatsache, dass Linda frei heraus ihre Meinung sagt, dadurch erklärt werden, dass sie aktive Feministin ist: Sie wird spätestens durch ihre Mitstreiterinnen gelernt haben, dass eine Frau mit ihren Ansichten keineswegs hinter dem Berg halten muss. Dasselbe gilt für ihr Single-Dasein. Warum ist Linda allein stehend? Weil, so könnte man mit BF argumentieren, Feministinnen typischerweise fürchten, in einer engen Partnerschaft wieder in die traditionelle Frauenrolle gedrängt zu werden.

Umgekehrt ermöglichen andere Teile der Daten eine Erklärung für die Feminismus-Komponente: Dass Linda sich schon als Studentin mit Diskriminierungsfragen beschäftigt hat, kann als eine der Ursachen für ihr aktuelles feministisches Engagement angesehen werden. Und schließlich bietet die Hypothese, dass Linda aktiv in der feministischen Bewegung ist, Aufschluss darüber, warum sie ausgerechnet als Bankangestellte arbeitet, obwohl sie dieser Beruf nach allem, was wir über sie wissen, schwerlich ausfüllen wird: Sie lebt ihre eigentlichen Interessen in ihren gesellschaftspolitischen Aktivitäten aus.

Bezieht man diese Faktoren in die Beurteilung der Annahmen ein, ist es alles andere als offensichtlich, dass die Konjunktion BF mindestens genauso unplausibel wie ihr Konjunkt B ist.⁹ Sie enthält mit B zwar einen Bestandteil, der *für sich* unglaublich ist; aber das bedeutet nicht, dass sie *als Ganzes* ebenso inakzeptabel ist. Es findet hier etwas statt, das man *epistemische Kompensation* nennen könnte: B lässt sich im Verbund mit F leichter schlucken – zum einen, weil das Zusammentreffen von Lindas Berufswahl und den Daten, die so gar nicht dazu passen wollen, durch F erklärlich wird; zum anderen, weil dann eine Reihe weiterer Erklärungsbeziehungen hergestellt werden können. BF ist aufgrund der geringeren Wahrscheinlichkeit vielleicht mit einem größeren Risiko verbunden als B allein, aber dem steht ein höheres Erklärungspotential gegenüber. Was bislang nicht harmonierte, lässt sich manchmal durch eine zusätzliche Annahme zu einem geschlosseneren Bild zusammensetzen, das mehr Sinn ergibt.

Hier ist ein zweites, weitaus einfacheres Exempel. Unser Datum sei diesmal, dass Linda wenig Freizeit hat; die Hypothesen seien B und BF. Auch dann scheint mir die Konjunktion besser als ihr Konjunkt dazustehen. Dass Linda in einer Bank arbeitet, bietet für sich noch keine Erklärung dafür, dass sie wenig freie Zeit hat, da sich Bankangestellte üblicherweise nach acht Stunden auf den Heimweg machen können. Nehmen

⁹ In Siebel 2003, 471, wird darauf verwiesen, dass Spohns Theorie der Rangfunktionen demselben Problem ausgesetzt ist.

wir aber auch noch an, dass Linda eine aktive Feministin ist, fällt es nicht schwer, das Datum zu erklären: Sie ist nach der Arbeit noch häufig mit ihren politischen Angelegenheiten beschäftigt. Der Probabilist ist dagegen darauf festgelegt, dass BF höchstens ebenso akzeptabel wie B ist. Er übersieht damit, dass Schlüsse auf die beste Erklärung einen in der Anreicherung von Hypothesen rechtfertigen können.

Zum Abschluss ein Beispiel aus der Wissenschaftsgeschichte. Nach der Entdeckung von Uranus bemerkte man, dass seine Umlaufbahn Anomalitäten aufwies: Aus den Fakten über die bislang bekannten Planeten und aus Newtons Gesetzen ergaben sich nicht die Positionen, die man beobachten konnte. Es wurde deswegen vermutet, dass ein weiterer Planet die Bahn beeinflusst. Nach einigen Berechnungen lag die Bewegung dieses Planeten fest; und tatsächlich fand man ihn – den Neptun – im September 1846 in der Nähe des vorhergesagten Ortes. Indem man den Newton'schen Gesetzen und den Kenntnissen über die bislang bekannten Planeten die Annahme eines weiteren Planeten hinzufügte, landete man hier bei einer logisch stärkeren Behauptung. Dennoch durfte sie eine höhere Plausibilität für sich verbuchen, weil sie eine Erklärung der Uranus-Bahn ermöglichte.

4. Eine Alternative: Kohärenz statt (nur) Wahrscheinlichkeit

Beispiele dieser Art machen deutlich, dass in die Beurteilung der Plausibilität einer Behauptung mehr einfließt als ihre Wahrscheinlichkeit. So ist z.B. ihr Erklärungspotential höchstgradig einschlägig. Ein Ansatz, der das berücksichtigt und deshalb mit solchen Fällen keinerlei Probleme hat, ist der *kohärenztheoretische*. Seine grundlegenden Ideen und seine Behandlung des Linda-Beispiels sollen hier kurz skizziert werden.

Der Kohärenztheoretiker geht davon aus, dass sich die Glaubwürdigkeit einer Hypothese daran bemisst, wie gut sie zu den Daten und dem weiteren Hintergrund passt. Die Passung – die Kohärenz – ist dabei im Wesentlichen von zwei Faktoren abhängig.¹⁰ Erstens ist zu konstatieren, dass *Inkonsistenzen* einen *negativen* Einfluss haben. Das gilt insbesondere für logische Widersprüche, in weniger starkem Ausmaß aber auch für schwächere Arten von Konflikten. Man denke an einen Schüler, der glaubt, er werde eine Prüfung bestehen, obwohl er sich der Tatsache bewusst ist, dass er sehr schlecht vorbereitet ist. Es ist zwar möglich, dass beide Annahmen wahr sind, aber sie passen nicht so recht zusammen. An dieser Stelle können natürlich auch Wahrscheinlichkeiten eine Rolle spielen. Die Wahrscheinlichkeit dafür, dass Tweety nicht fliegen kann, gegeben dass es sich um einen Vogel handelt, ist relativ gering. Diese Hypothesen können deshalb als in einem schwachen Sinn inkonsistent aufgefasst werden, was ihrer Kohärenz abträglich ist.

¹⁰ Vgl. insbesondere Bartelborth 1996, Kap. IV.D, IV.F.3; BonJour 1985, Kap. 5.3; Thagard 1992, Kap. 4; und für den Einfluss von Konflikten auch Lehrer 1990, 148. – Ein dritter Punkt, der häufig erwähnt wird, ist, dass isolierte Subsysteme die Kohärenz schmälern. Dieser Faktor ist für das Linda-Beispiel allerdings so gut wie irrelevant.

Zweitens stehen auf der *postiven* Seite *Folgerungs-* und *Erklärungsbeziehungen*. Dass ein Aussagensystem keine Inkonsistenzen aufweist, impliziert ja noch nicht, dass es kohärent ist. Wenn seine Elemente aber in inferentiellen oder explanatorischen Beziehungen stehen, dann kann sich auch Kohärenz ausbreiten. Je mehr Annahmen sich z.B. durch andere Annahmen erklären lassen und je besser die Erklärungen sind, umso geschlossener ist das Ganze. Wenn etwa H_1 D_1 und D_2 erklärt, während H_2 nur Aufschluss über D_1 gibt, dann ist $\{H_1, D_1, D_2\}$ kohärenter als $\{H_2, D_1, D_2\}$.

Wie steht es im Rahmen dieses Ansatzes um das Linda-Beispiel von Tversky und Kahneman? Am Rande erwähnt sei, dass Paul Thagard (1989; 1992, Kap. 4) und Daniel Schoch (2000; 2002) die Ideen der Kohärenztheorie in Computermodelle gegossen haben, die quantitative Kohärenzabschätzungen liefern. Speist man in Thagards Kohärenzprogramm ECHO¹¹ die entsprechenden Informationen ein, dann kommt es zu dem gewünschten Resultat: Die Geschichte über Linda passt mit der Konjunktion BF *besser* zusammen als mit B allein. Es lässt sich aber auch ohne jeden technischen Aufwand deutlich machen, warum das so ist.

Zwischen der Beschreibung von Linda und der Hypothese, dass sie eine Bankangestellte ist, lassen sich keine Folgerungs- oder Erklärungsbeziehungen herstellen. Stattdessen stößt man auf eine Reihe von (nicht-logischen) Inkonsistenzen: Dass z.B. eine ehemalige Philosophiestudentin ausgerechnet Bankangestellte wird, kommt nicht gerade häufig vor. Zusammen mit dem ersten Punkt führt das dazu, dass wir es mit einem inkohärenten Ganzen zu tun haben.

Anders sieht es aus, wenn wir die konjunktive Hypothese betrachten, in der Linda zusätzlich feministische Aktivitäten zugeschrieben werden. Es schleicht sich dann zwar eine weitere Inkonsistenz ein, weil die Hypothese selbst aus zwei Bestandteilen besteht, die miteinander konfliktieren. Dafür ergeben sich durch die Feminismus-Komponente aber vielfältige Erklärungsmöglichkeiten, die für eine Kompensation der negativen Faktoren sorgen. Wie im letzten Abschnitt gezeigt, lässt die Feminismus-Annahme sich durch Teile der Daten erklären, so wie auch umgekehrt einige der Daten auf ihrer Grundlage erklärbar werden. Im Übrigen würde sie verständlich machen, wie Linda mit dem Konflikt umgeht, der zwischen ihrem Beruf und den Daten besteht. Es fließen also eine Reihe von kohärenzstiftenden Beziehungen ein, so dass – alles in allem – mehr Kohärenz zu attestieren ist.

Noch besser ist die Situation, wenn man nur die Hypothese nimmt, dass Linda eine aktive Feministin ist. Hier lassen sich wie zuvor viele Erklärungsbeziehungen herstellen, während zugleich die Inkonsistenzen wegfallen, die aufgrund des Bankangestellten-Konjunks entstanden sind. Die Kohärenz ist deshalb noch größer als im Falle der Konjunktion. Der kohärenztheoretische Ansatz führt somit zu der Einschätzung, die wir haben wollen: Die Feminismus-Hypothese ist einleuchtender als die Konjunktion – und die wiederum plausibler als die Bankangestellten-Hypothese.

Daraus schließe ich, dass die Kohärenztheorie dem Probabilismus in einem wichtigen Punkt überlegen ist: Sie lässt zu, dass eine Hypothese H_1 hinsichtlich ihrer Plausibilität

¹¹ Eine Java-Version findet sich unter <http://cogsci.uwaterloo.ca/JavaECHO/jecho.html>.

besser dasteht als eine Aussage H_2 , obwohl H_2 eine logische Konsequenz von H_1 ist. Wenn durch H_1 mehr Erklärungen möglich werden, dann kann damit das Manko aufge-
wogen werden, dass H_1 weniger wahrscheinlich ist.

5. Eine psychologische Hypothese und ihre normativen Konsequenzen

Diese Überlegungen werfen dann auch ein neues Licht auf die ursprünglichen Experimente von Tversky und Kahneman. Sie haben ihre Versuchspersonen nicht nach der Plausibilität, sondern nach der *Wahrscheinlichkeit* der Annahmen über Linda gefragt. Bei nur zwei Möglichkeiten – „Bankangestellte“ und „Bankangestellte und aktive Feministin“ – haben 90 % der Teilnehmer ohne Statistikausbildung die Konjunktion als „wahrscheinlicher“ hervorgehoben. Unter denen mit Statistikausbildung lag dieser Anteil zwar nur bei 50 %; aber als diese beiden Annahmen in eine Liste mit insgesamt acht Alternativen eingebettet wurden und eine Rangfolge erstellt werden sollte, bewerteten schließlich auch über 80 % der Personen mit Statistikkenntnissen die Konjunktion höher.

Die Bevorzugung der Konjunktion wird im Allgemeinen als *Fehler* betrachtet – und letztlich als ein Hinweis auf eine gewisse Irrationalität. Aber das scheint mir etwas voreilig zu sein. Natürlich, wenn man das Wörtchen „wahrscheinlich“ im Sinne des mathematischen Kalküls interpretiert, dann kann eine Konjunktion nicht wahrscheinlicher als ihre Konjunkte sein. Gerd Gigerenzer (u.a. 2001, 95f.) und seine Mitarbeiter haben aber darauf hingewiesen, dass die Versuchspersonen diesen Ausdruck auch anders verstehen können, da er mehrdeutig ist. Neben seiner speziellen technischen Bedeutung hat er noch eine Reihe sehr viel alltäglicherer Bedeutungen – wie eben „glaubhaft“ oder „plausibel“. Wenn die Aufgabenstellung in diese Richtung gedeutet wird, dann wird mit der Wahl der Konjunktion aber gar nicht behauptet, dass sie *im technischen Sinn* wahrscheinlicher ist.

Hertwig und Gigerenzer (1999, 278) machen außerdem deutlich, dass es selbst dann, wenn einem der Wahrscheinlichkeitskalkül bekannt ist, einen guten Grund gibt, die gestellte Aufgabe anders auszulegen: In der mathematischen Lesart wäre Paul Grices *Relevanzmaxime* verletzt (vgl. Grice 1975, 27). Diese Maxime ist eine Konversationsregel, von der üblicherweise angenommen wird, dass die Gesprächsteilnehmer sie befolgen. Sie sagt, ganz grob, dass jeder Beitrag relevant für das Thema und das Ziel des Dialogs ist. Auf den vorliegenden Fall angewandt würde das etwa heißen, dass alle Teile der Aufgabenstellung von Belang sind. Nimmt man aber die technische Interpretation von „wahrscheinlich“, so ist die Kurzcharakterisierung von Linda vollkommen irrelevant für die Frage, die es zu beantworten gilt. Denn die Wahrscheinlichkeit einer Konjunktion, gegeben eine Annahme, kann nicht höher sein als die Wahrscheinlichkeit von einem der Konjunkte, gegeben dieselbe Annahme – *ganz egal, wie die Annahme lautet*. Ob nun von Linda berichtet wird, dass sie Philosophie studiert hat oder dass sie eine Banklehre hinter sich hat, für *jede* Beschreibung D gilt: $P(BF|D) \leq P(B|D)$. Es bietet sich deshalb an, den Ausdruck „wahrscheinlich“ nicht im Sinne des Kalküls zu ver-

stehen. Ansonsten müssten die zur Verfügung gestellten Daten nämlich als bedeutungslos betrachtet werden.¹²

Und tatsächlich ist die von Tversky und Kahneman intendierte Lesart der Aufgabenformulierung ja nicht weit von einer Scherzfrage entfernt. Wer rechnet als Teilnehmer an einem ernsthaften psychologischen Experiment schon damit, dass die einzigen konkreten Informationen, die einem vorgelegt werden, überflüssig sind. Tversky und Kahneman müssen sich deshalb den Vorwurf gefallen lassen, dass sie ihre Versuchspersonen zuerst auf eine Weise in die Irre führen, die nur schwer zu durchschauen ist, um sie anschließend der Irrationalität zu bezichtigen.

Ganz andere Ergebnisse bekommt man dann auch, wenn die Aufgabenstellung die Wahrscheinlichkeitstheoretische Ausrichtung wirklich deutlich macht. So hat Fiedler (1998) die ursprüngliche Frage von Tversky und Kahneman in eine Frage nach Häufigkeiten umformuliert: „Es gibt 100 Personen, auf die die gegebene Beschreibung zutrifft. Wie viele von ihnen werden (a) Bankangestellte und (b) Bankangestellte und aktiv in der feministischen Bewegung sein?“ Wer (b) hier eine höhere Zahl als (a) zuordnet, der begeht in der Tat einen Fehler. Aber das wurde auch nur noch von 22 % der Teilnehmer gemacht.

Bleiben wir bei den Experimenten von Tversky und Kahneman, und gehen wir mit Gigerenzer davon aus, dass bei vielen der Versuchspersonen gar keine probabilistischen Überlegungen im Hintergrund stehen, weil sie den Ausdruck „wahrscheinlich“ nicht im Sinne des mathematischen Kalküls auslegen. Was könnte sie dann dazu veranlassen, der Konjunktion eine größere „Wahrscheinlichkeit“ zuzugestehen? Mein Vorschlag dürfte nicht überraschen: Sie versuchen, Kohärenz zu maximieren. Sie bewerten diejenige Annahme als „wahrscheinlicher“, die zu einem stimmigeren Bild von Linda führt.

Mit dieser Behauptung bewege ich mich aus dem Gebiet der Erkenntnistheorie hinaus, indem ich eine *psychologische* Hypothese aufstelle. Sie gilt es erst einmal experimentell zu überprüfen; und da Tversky und Kahneman (u.a. 1982, 97) einen eigenen Erklärungsvorschlag ins Spiel gebracht haben, wären die Versuche so zu gestalten, dass auf der Basis ihrer Erklärungsalternative andere Ergebnisse zu erwarten sind als auf der Grundlage meiner.¹³ Es kann an dieser Stelle aber zumindest darauf verwiesen werden, dass mein Vorschlag nicht ad hoc ist. Thagard hat dafür argumentiert, dass Kohärenz u.a. in der Begriffsformation, dem Handlungsverstehen und der Bildung von Emotionen eine

¹² Vgl. auch Gigerenzer 2001, 96. Adler (1991, 255) und andere haben betont, dass die Relevanzmaxime die Versuchspersonen auch dazu bringen kann, „Linda ist eine Bankangestellte“ so zu verstehen, dass damit behauptet wird, Linda sei *keine* aktive Feministin. Einige Experimentatoren haben diese Implikatur abgeblockt, indem sie Formulierungen eingesetzt haben wie „Linda ist eine Bankangestellte, ob sie nun aktiv in der feministischen Bewegung ist oder nicht“. Da in den meisten Studien immer noch 50 % der Probanden die Konjunktion bevorzugten, scheint Adlers Erklärung noch nicht der Weisheit letzter Schluss zu sein: Die oben beschriebene Art, die Relevanzmaxime einzuhalten, ist weiter im Spiel. Eine Übersicht über die entsprechenden Experimente bieten Hertwig und Gigerenzer (1999, 297f.).

¹³ Die Toulouser Psychologen Jean François Bonnefon und Rui Da Costa Neves wollen sich entsprechender Experimente annehmen.

gewichtige Rolle spielt.¹⁴ Wer die Reaktion der Probanden im Linda-Experiment auf Kohärenzmaximierung zurückführt, verweist also nur auf einen Faktor, der im menschlichen Denken sowieso schon seinen Platz hat.

Wenn diese Idee richtig ist, dann ist der Irrationalitätsvorwurf alles andere als berechtigt. Erstens liegt es aufgrund der Relevanzmaxime nahe, die Aufgabenstellung so zu deuten, dass es nicht um Wahrscheinlichkeit in der technischen Bedeutung geht. Stanovich und West (2000, 655; meine Übers.) weisen darauf hin, dass man auch dann von Irrationalität reden kann, wenn die Interpretation der Versuchspersonen „so bizarr [ist] – so weit von dem entfernt, was die Worte in den Instruktionen sagen –, dass es sich um einen ernsthaften kognitiven Fehler handelt“. Aber dieses Verdikt ist hier nicht angebracht. Zweitens zeugt die Bevorzugung der Konjunktion – wenn sie auf die höhere Kohärenz abzielt, die sich so ergibt – von einem Verständnis von Plausibilität, das nicht so eingeschränkt ist wie dasjenige des Probabilisten. Wie schon gesagt: Die Konjunktion steht zwar bezüglich ihrer Wahrscheinlichkeit schlechter da, aber sie ist deswegen nicht unplausibler. Wenn man bedenkt, welches Erklärungspotential durch die Feminismus-Komponente eröffnet wird, ist es keineswegs irrational, der risikoreicheren Hypothese den Vorrang zu geben.

Gigerenzer weist in seinen Texten gelegentlich auf den Paläontologen Stephen Gould hin, der in *Bully for Brontosaurus* (1992, 469) schreibt: „I know that the [conjunction] is less probable, yet a little homunculus in my head continues to jump up and down, shouting at me – ,but she can't just be a bankteller; read the description.““ Wie Gigerenzer (1994, 131) bin ich der Meinung, dass Gould das Gezeter seines Homunkulus nicht leichtfertig abtun sollte. Denn er könnte ein kohärenter Denker sein, der nicht nur Wahrscheinlichkeiten im Sinn hat, sondern bei der Beurteilung von Hypothesen weitere essentielle Faktoren wie ihre Erklärungsstärke einbezieht.¹⁵

Literatur

- Adler, J. (1991): „An Optimist's Pessimism: Conversation and Conjunction“, in: *Poznan Studies in the Philosophy of the Sciences and the Humanities* 21, 251–282.
- Bartelborth, T. (1996): *Begründungsstrategien. Ein Weg durch die analytische Erkenntnistheorie*. Berlin: Akademie Verlag.
- BonJour, L. (1985): *The Structure of Empirical Knowledge*. Cambridge/M. und London: Harvard University Press.
- Bovens, L. und S. Hartmann (2003): *Bayesian Epistemology*. New York und Oxford: Oxford University Press.

¹⁴ Thagard 1997; Thagard und Kunda 1998; Thagard 2000, Kap. 6. Zur Rolle von Erklärungen in der menschlichen Kognition siehe die Beiträge in Keil und Wilson 2000.

¹⁵ Dieser Beitrag entstand im Rahmen des DFG-Projektes *Erklärungskohärenz*. Ich danke Nick Allott, Thomas Bartelborth, Jean François Bonnefon, David Chart, Rui Da Costa Neves, Christoph Dörge, Ralph Hertwig, Paul Hoyningen-Huene, Daniel Schoch, Wolfgang Spohn und Mark Textor für die Diskussion von Vorläufern des Textes oder entsprechender Vorträge.

- Carnap, R. (1962): *Logical Foundations of Probability*. (2. Aufl.) Chicago: University of Chicago Press.
- Earman, J. (1992): *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*. Cambridge/M.: MIT Press.
- Eells, E. und B. Fitelson (2002): „Symmetries and Asymmetries in Evidential Support“, in: *Philosophical Studies* 107, 129–142.
- Fiedler, K. (1988): „The Dependence of the Conjunction Fallacy on Subtle Linguistic Factors“, in: *Psychological Research* 50, 123–129.
- Fitelson, B. (1999): „The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity“, in: *Philosophy of Science* 66 (Proceedings), S362–S378.
- (2001): *Studies in Bayesian Confirmation Theory*. Dissertation, University of Wisconsin at Madison. Online: <http://fitelson.org/thesis.pdf>.
- (2003): „A Probabilistic Theory of Coherence“, in: *Analysis* 63, 194–199.
- Gigerenzer, G. (1994): „Why the Distinction between Single-Event Probabilities and Frequencies is Important for Psychology (and vice versa)“, in: G. Wright und P. Ayton (Hrsg.): *Subjective Probability*. Chichester: Wiley, 129–161.
- (2001): „Content-Blind Norms, no Norms, or Good Norms? A Reply to Vranas“, in: *Cognition* 81, 93–103.
- Glymour, C. (1998): „Why I am not a Bayesian“, in: M. Curd und J. A. Cover (Hrsg.): *Philosophy of Science. The Central Issues*. New York und London: Norton, 584–606.
- Gould, S. J. (1992): *Bully for Brontosaurus. Further Reflections in Natural History*. London: Penguin Books.
- Grice, H. P. (1975): „Logic and Conversation“, in: *Studies in the Way of Words*. Cambridge/M. und London 1989: Harvard University Press, 22–40. Urspr. veröff. in: P. Cole und J. Morgan (Hrsg.): *Syntax and Semantics 3: Speech Acts*. New York: Academic Press, 41–58.
- Hertwig, R. und G. Gigerenzer (1999): „The ‚Conjunction Fallacy‘ Revisited: How Intelligent Inferences look like Reasoning Errors“, in: *Journal of Behavioral Decision Making* 12, 275–305.
- Howson, C. und P. Urbach (1993): *Scientific Reasoning. The Bayesian Approach*. (2. Aufl.) Chicago und Lasalle/Ill.: Open Court.
- Jaynes, E. T. (1996): *Probability Theory: The Logic of Science*. Manuskript. Online: [//www.inrialpes.fr/sharp/pub/laplace/Jaynes/prob.html](http://www.inrialpes.fr/sharp/pub/laplace/Jaynes/prob.html).
- Kahneman, D. und A. Tversky (1982): „On the Study of Statistical Intuitions“, in: D. Kahneman, P. Slovic und A. Tversky (Hrsg.): *Judgment under Uncertainty. Heuristics and Biases*. Cambridge: Cambridge University Press, 493–508. Urspr. veröff. in: *Cognition* 11, 123–141.
- Keil, F. C. und R. A. Wilson (Hrsg.) (2000): *Explanation and Cognition*. Cambridge/M. und London: MIT Press.
- Lehrer, K. 1990: *Theory of Knowledge*. London und New York: Routledge.

- Mackie, J. L. (1969): „The Relevance Criterion of Confirmation“, in: *The British Journal for the Philosophy of Science* 20, 27–40.
- Olsson, E. (2002): „What is the Problem of Coherence and Truth?“, in: *The Journal of Philosophy* 94, 246–272.
- Schoch, D. (2000): „Explanatory Coherence“, in: *Synthese* 122, 291–311.
- (2002): „Computationale Abduktion“, in: M. Siebel (Hrsg.): *Kommunikatives Verstehen*. Leipzig: Leipziger Universitätsverlag, 198–219.
- Shogenji, T. (1999): „Is Coherence Truth Conducive?“, in: *Analysis* 59, 338–345.
- Siebel, M. (2003): Rezension von *Erkenntnistheorie. Positionen zwischen Tradition und Gegenwart* (hrsg. v. T. Grundmann, Mentis: Paderborn 2001), in: *Zeitschrift für philosophische Forschung* 57, 467–471.
- (2004a): „On Fitelson’s Measure of Coherence“, in: *Analysis* 64, 189–190.
- (2004b): „Der Rabe und der Bayesianist“, erscheint in: *Zeitschrift für allgemeine Wissenschaftstheorie* 35.
- (2004c): „Confirmation and Coherence“, Manuskript, im Begutachtungsprozess bei *Synthese*.
- (2005): „Probabilistic Measures of Coherence“, erscheint in *Erkenntnis*.
- Sober, E. (1994): „No Model, no Inference: A Bayesian Primer on the Grue Problem“, in: D. Stalker (Hrsg.): *Grue! The New Riddle of Induction*. Chicago: Open Court, 225–240.
- Stanovich, K. E. und R. F. West (2000): „Individual Differences in Reasoning: Implications for the Rationality Debate?“, in: *Behavioral and Brain Sciences* 23, 645–665.
- Thagard, P. (1989): „Explanatory Coherence“, in: *Behavioral and Brain Sciences* 12, 435–502.
- (1992): *Conceptual Revolutions*. Princeton: Princeton University Press.
- (1997): „Coherent and Creative Conceptual Combinations“, in: T. B. Ward, S. M. Smith und J. Viad (Hrsg.): *Creative Thought: An Investigation of Conceptual Structures and Processes*. Washington, D.C.: American Psychological Association, 129–141.
- (2000): *Coherence in Thought and Action*. Cambridge/M. und London: MIT Press.
- Thagard, P. und Z. Kunda (1998): „Making Sense of People: Coherence Mechanisms“, in: S. J. Read und L. C. Miller (Hrsg.): *Connectionist Models of Social Reasoning and Social Behavior*. Hillsdale/N.J.: Erlbaum, 3–26.
- Tversky, A. und D. Kahneman (1982): „Judgements of and by Representativeness“, in: D. Kahneman, P. Slovic und A. Tversky (Hrsg.): *Judgment under Uncertainty. Heuristics and Biases*. Cambridge: Cambridge University Press, 84–98.