

Frequency-Validity-Effekt und Hindsight-Bias: Unterschiedliche Phänomene – gleiche Prozesse?¹

Ralph Hertwig

'Persil wäscht so weiß, weißer geht's nicht.' 'Scheidung im Hause Windsor – Prinzessin Diana und Prinz Charles trennen sich.' 'Die CDU bleibt stärkste Kraft im Landtag.' Aussagen und Slogáns áhnlich diesen prasseln úberall und ununterbrochen auf uns nieder, und oftmals sind wir gezwungen, uns úber die Wahrheit oder Unwahrheit dieser Aussagen eine Meinung zu bilden, obgleich uns die Informationen, die eine sichere Beurteilung ermúglichen wúrdien, gegenwártig nicht zur Verfügung stehen. In diesen, durch Mangel an Information gekennzeichneten Situationen beziehen wir entweder einen Indifferenz-Standpunkt, d.h. wir halten Wahrheit und Unwahrheit der Aussage für gleichwahrscheinlich, oder aber wir fállen auf der Grundlage der unmittelbar zugánglichen Information eine Entscheidung unter Vorbehalt. Der Vorbehalt kommt darin zum Ausdruck, daß das Urteil für oder gegen die Richtigkeit der Aussage vorláufig ist und für die Urteile keine absolute Sicherheit in Anspruch genommen wird: Man hált eine Aussage für wahrscheinlich wahr (oder falsch), ist sich aber nicht absolut sicher. Sollten wir zu einem spáteren Zeitpunkt erneut mit einer bereits gehörten Aussage konfrontiert werden, dann kúnnnte man, sofern sich an unserer Wissensbasis úber den in der Aussage angesprochenen Sachverhalt nichts geándert hat, erwarten, daß wir unsere frühere Einschätzung identisch reproduzieren. Als erste untersuchten Hasher, Goldstein und Toppino (1977) den Einfluß der Wiederholung derselben Aussagen auf das Urteilsverhalten unter experimentell kontrollierten Bedingungen und fanden, daß „... *the repetition of a plausible statement increases a person's belief in the referential validity or truth of that statement*”² (S. 111). Dies bezeichneten sie als *Frequency-Validity-Effekt*.

¹ Ich móchte an dieser Stelle Regina Bernhaupt, Hartmut Blank, Edgar Erdfelder, Gerd Gigerenzer, Wolfgang Hell, Günther Kebeck, Toni Kühberger und Peter Sedlmeier für viele kritische Anregungen und Hinweise danken. Besonderen Dank aber schulde ich Ulrich Hoffrage. Diese Arbeit ist Ausdruck unserer zahllosen Diskussionen, in denen wir unsere Ideen sortierten, bewerteten, oft verwarfen und sie manchmal auch für gut befanden. Weiterhin danke ich dem Fonds zur Förderung der wissenschaftlichen Forschung für finanzielle Unterstützung (FWF-Projekt Nr. P8842-MED).

² An Stelle der Formulierung „person's belief in the referential validity or truth of that statement” werde ich den Begriff „Konfidenz-Urteil” verwenden.

Hätte man vor den letzten Wahlen zum Bundestag die Wahlberechtigten mehrere Male um eine Einschätzung der Aussage *Die CDU/CSU/FDP-Koalition gewinnt wiederum die absolute Mehrheit der Mandate* gebeten, wäre nach Hasher et al. (1977) zu erwarten gewesen, daß der Glaube an die Richtigkeit dieser Aussage mit der Wiederholung des Urteils zunimmt. Hätte man hingegen nach der ersten Einschätzung die Bundestagswahl 1990 abgewartet und die Personen nach Bekanntgabe des amtlichen Endergebnisses gebeten, sich ihres ursprünglichen Urteils zu erinnern, wäre in den Erinnerungen ein anderer Effekt zutage getreten, der in der Literatur als *Hindsight-Bias* (im Deutschen gelegentlich auch als „Rückschau-Fehler“) bezeichnet wird. Der „Fehler“ hätte darin bestanden, daß die befragten Personen sich in ihren erinnerten Urteilen näher an dem tatsächlichen Wahlausgang glaubten, als sie es ursprünglich waren: Ihre vormalige subjektive Sicherheit in die Richtigkeit der Aussage *Die CDU/CSU/FDP-Koalition gewinnt wiederum die absolute Mehrheit der Mandate* wäre in der Erinnerung überschätzt worden. Hätte hingegen die Koalition die absolute Mehrheit verloren, wäre ihre vormalige Sicherheit unterschätzt worden. Fischhoff (1977, S. 349) charakterisierte den Hindsight-Bias mit den Worten: „... (a) that telling people that an event has occurred increases their subjective probability that it was going to happen and (b) that people underestimate the effect that hearing such reports has on their perceptions.“

Eine Bezugnahme auf die Forschung zum jeweils anderen Effekt fand weder in den Arbeiten zum Hindsight-Bias noch zum Frequency-Validity-Effekt statt. Dies ist erstaunlich, da zum ersten Urteilszeitpunkt einer Frequency-Validity- und einer Hindsight-Bias-Untersuchung völlig identische Anforderungen an die Versuchspersonen gestellt werden. Erst im Anschluß an das erste Urteil werden Unterschiede in den Anordnungen offenkundig. Wo im Frequency-Validity-Design ein weiteres Konfidenz-Urteil zum gleichen Ereignis erbeten wird, erhalten die Versuchspersonen im Hindsight-Bias-Design eine Information über das Ereignis und werden anschließend gebeten, sich ihres ersten Urteils zu erinnern.

Ziel dieses Beitrages ist die Skizzierung eines Modells, mit dem die Grundlage geschaffen wird, sowohl den Frequency-Validity-Effekt wie auch den Hindsight-Bias innerhalb des gleichen theoretischen Rahmens zu erklären. Ausgangspunkt ist die Annahme, daß sowohl anläßlich der wiederholten Konfidenz-Urteile (Frequency-Validity-Design) wie auch anläßlich der Erinnerungs-Urteile (Hindsight-Bias-Design) Konstruktions-, Rekonstruktions- und Erinnerungsprozesse stattfinden. Somit sind in diesem Modell beide Effekte Ausdruck prinzipiell gleicher Prozesse. Unterschiede zwischen den Effekten bestehen dann darin, wieviel Prozent der Urteile durch die jeweiligen Prozesse aufgeklärt werden können.

Zunächst werden die wesentlichen Methoden, Ergebnisse und theoretischen Erklärungen zum Frequency-Validity-Effekt skizziert. Die Forschung zum Hindsight-Bias wird nur insoweit vorgestellt, wie sie für das Verständnis des Modells notwendig erscheint. Der größte Teil dieses Beitrages wird der

Einführung des Modells gewidmet. Anschließend werden die Vorhersagen, die aus dem Modell folgen, formuliert. Beschlossen wird dieser Beitrag mit einem Blick auf noch offene Fragen.

Frequency-Validity-Effekt: Methoden, Ergebnisse und Erklärungsansätze

Methoden und Ergebnisse

Die Beziehung zwischen bloßer Wiederholung und zunehmender Sicherheit in die „Wahrheit“ einer wiederholten Aussage untersuchten Hasher et al. (1977), indem sie ihren Probanden 60 plausible Aussagen aus 10 Wissensbereichen darboten. Diese Aussagen (30 wahre und 30 falsche) waren so ausgewählt worden, daß die Vpn sich über deren Wahrheit oder Unwahrheit unsicher sein sollten (z.B.: *Die Volksrepublik China wurde 1947 gegründet*). Die Probanden beurteilten ihre subjektive Sicherheit in die Entscheidung für einen Wahrheitswert (wahr/falsch) auf einer 7-Punkte-Skala (numerisch und verbal verankert, z.B.: 1/7= 'definitely false/true'; 4= 'uncertain'). Für einen Teil der Aussagen (Experimental-Aussagen) wurden in einer zweiten und dritten Sitzung jeweils zwei Wochen später – im Kontext neuer, bislang noch nicht dargebotener Aussagen (Kontroll-Aussagen) – erneut Konfidenz-Urteile eingefordert. Hasher et al. (1977) konstatierten für die Experimental-Aussagen – ungeachtet ihres Wahrheitswertes – eine Zunahme des gemittelten Urteils der subjektiven Sicherheit.³ An diesem experimentellen Design wurden in späteren Arbeiten Modifikationen vorgenommen, die dadurch motiviert waren, Variablen bzw. Bedingungen zu untersuchen, von denen vermutet wurde, daß sie Auftreten und Ausmaß der Zunahme des Konfidenz-Urteils beeinflussen könnten. Im folgenden werden die Ergebnisse für vier Variablen – *Zeitabstand zwischen den Darbietungen, Vertrautheit mit dem Gegenstand der Aussage, präjudizierende Lernbedingungen und Wiedererkennung* – herausgegriffen.

Zeitabstand zwischen den Darbietungen

Unabhängig davon, ob die Darbietungen der Aussagen durch wenige Minuten (Schwartz, 1982) oder Wochen (Bacon, 1979) getrennt wurden, manifestierte sich der Frequency-Validity-Effekt. Gigerenzer (1984) erhob bei verschiedenen Versuchspersonengruppen die Urteile nach einem ein- oder zweiwöchigen Zeitintervall und konnte auch bei diesem direkten Vergleich der Zeitintervalle

³ Berechnet man die Effektgröße nach Cohen (1988) ergeben sich für die Zunahmen der Konfidenz-Urteile der Experimental-Aussagen von der 1. zur 2. und von der 2. zur 3. Sitzung folgende Effektgrößen: a) Wahre Aussagen: $d'_{(1,2)} = 0,51$; $d'_{(2,3)} = 0$ b) Falsche Aussagen: $d'_{(1,2)} = 0,67$; $d'_{(2,3)} = 0,25$. Die Effektgrößen von der 1. zur 2. Sitzung sind also deutlich größer als von der 2. zur 3. Sitzung.

keine Unterschiede in der Zunahme der Konfidenz-Urteile (d.h. im Anstieg der über alle Vpn gemittelten Konfidenz-Urteile) für die Experimental-Aussagen feststellen.

Vertrautheit mit dem Gegenstand der Aussage

Die präexperimentelle Vertrautheit mit Konzepten, die in den Aussagen angesprochen werden, ist die zentrale Variable in der von Schwartz (1982) formulierten Substitutions-Hypothese. Beinhaltet eine Aussage präexperimentell erlernte Konzepte (z.B.: *Ernest Hemingway, 1. Weltkrieg*), dann evoziert diese Aussage eine Vertrautheitserfahrung. Dies geschieht selbst dann, wenn die Aussage in toto für die Person völlig neu ist (z.B.: *Ernest Hemingway war im 1. Weltkrieg Fahrer eines Krankenwagens*). Sollte nun die verfügbare Information im Gedächtnis nicht ausreichen, um den Wahrheitswert der Aussage zu bestimmen, wird an Stelle des Konfidenz-Urteils ein Urteil über die präexperimentelle Vertrautheit gesetzt. Die Höhe der subjektiven Sicherheit in die Richtigkeit der Aussage richtet sich dann also nach dem Ausmaß der Vertrautheit, das die Aussage erweckt. Fände bei der Wiederholung der Aussagen im experimentellen Kontext ein zumindest teilweiser Verlust der Diskriminationsfähigkeit zwischen präexperimentell und experimentell erworbener Vertrautheit statt, könnte die wiederholte Darbietung zu einer illusionären Erhöhung der präexperimentellen Vertrautheit und folglich zu einem höheren Konfidenz-Urteil führen. Schwartz überprüfte diese Hypothese, indem sie nach dreimaliger Darbietung der Aussagen in einer Gruppe ein Urteil der präexperimentellen Vertrautheit, in einer anderen ein Konfidenz-Urteil auf derselben Skala erfragte. Obgleich die Darbietungen im experimentellen Kontext zu einer Erhöhung beider Urteile für die Experimental-Aussagen führten, konnte die Substitutions-Hypothese in dieser strengen Version nicht bestätigt werden. Im Gegensatz zur Vorhersage waren die Verteilungen der Urteile über den Skalen nicht identisch, die Urteile zur präexperimentellen Vertrautheit waren im Mittel niedriger als die Konfidenz-Urteile, und erstere unterschieden sich in Abhängigkeit vom tatsächlichen Wahrheitswert der Aussagen.

Präjudizierende Lernbedingungen

Der Frage, welche Auswirkungen präjudizierende Lernbedingungen haben, wandten sich Begg, Armour und Kerr (1985) zu, indem sie Aussagen zunächst mit affirmativen oder diskreditierenden Vordersätzen versahen und darboten (z.B.: *Es wird gemeinhin geglaubt, daß ... versus Wenige Leute glauben, daß ...*). Anschließend wurden die Aussagen ohne Vordersatz erneut präsentiert, und es wurde erstmalig ein Konfidenz-Urteil erhoben. Erstaunlicherweise erhielten sowohl die vormalig positiv wie auch die negativ beleumundeten Aussagen im Mittel höhere Urteile als neue, nicht-wiederholte Aussagen. Innerhalb der wiederholten Aussagen erhielten allerdings die ehemals mit

positiven Vordersätzen verknüpften Aussagen höhere Urteile als die in ihrer Glaubwürdigkeit beeinträchtigten. Diese Rangfolge führte die Autoren zu der Vermutung, daß die Urteile der Personen im wesentlichen durch die inhaltlichen Details der Aussagen und nicht durch die präjudizierenden Vordersätze beeinflusst würden. Der Effekt der Vordersätze bestehe lediglich darin, die Erinnerungswahrscheinlichkeit an die Details der Aussagen zu beeinflussen: In einem diskreditierenden Kontext erlernte Details werden weniger gut erinnert.

Wiedererkennung

Die experimentelle Untersuchung der bisher vorgestellten und weiterer ungenannter Variablen (z.B.: visuelle vs. auditorische Darbietung der Aussagen, Geschlecht des Sprechers) war weniger durch eine Theorie als durch das Bemühen, die Robustheit des Phänomens zu testen, geleitet. Weniger pragmatisch als vielmehr theoretisch war das Interesse an der letzten vorzustellenden Variablen. Die Autorengruppe Bacon (1979) untersuchte die Frage, ob die entscheidende Variable, die zu einer Erhöhung des Konfidenz-Urteils für wiederholte Aussagen führe, tatsächlich wie Hasher et al. (1977) postulierten die Repetition per se darstellt oder ob nicht die Erinnerung, daß eine Aussage bereits früher schon dargeboten wurde, die Grundlage des Effekts sein könnte. Konkret operationalisiert bedeutete dies: Die Erinnerung in der zweiten Sitzung, daß eine Aussage bereits in der ersten Sitzung dargeboten wurde, geht mit einem – im Vergleich zu den als nicht-wiederholt eingeschätzten Aussagen – höheren Konfidenz-Urteil einher. Die kritische Variable, die den Frequency-Validity-Effekt bedingt, ist also nach Bacon (1979) nicht der faktische (wiederholt vs. nicht-wiederholt) sondern der subjektive, zugeschriebene Status („alt“, „neu“) der Aussage. Der letztere kann in einem Wiedererkennungs-Urteil erfragt werden. Bacon führte ein Experiment durch, das eine Beurteilung der Erklärungskraft der beiden unabhängigen Variablen – Repetition per se versus subjektives Wiedererkennungs-Urteil – möglich machen sollte. In Ergänzung zu Hasher et al. (1977) wurde in der zweiten Sitzung neben dem Konfidenz-Urteil ein Wiedererkennungs-Urteil erhoben. Die wesentlichen Vorhersagen, die einer Prüfung unterzogen wurden, waren: Erstens, das gemittelte Konfidenz-Urteil über den Wahrheitswert der Aussagen, die in der zweiten Sitzung fälschlicherweise als wiederholt erinnert werden, sollte sich – nach Bacon – nicht von dem Urteil über jene Aussagen unterscheiden, die korrekt als wiederholt erinnert werden. Zweitens, das gemittelte Konfidenz-Urteil über den Wahrheitswert von Aussagen, die fälschlicherweise als nicht-wiederholt erinnert werden, sollte niedriger sein als das Konfidenz-Urteil über (korrekt oder fälschlicherweise) als wiederholt erinnerte Aussagen. Die Ergebnisse, die Bacon (1979) berichtet, bestätigen diese Vorhersagen, die im klaren Widerspruch zu der Annahme stehen, bloße Wiederholung der Aussagen reiche hin, um den Frequency-Validity-Effekt zu erzielen. Arkes, Hackett und Boehm

(1989, Experiment 1A) haben diese Vorhersagen von Bacon (1979) gleichfalls bestätigt.

Die Untersuchung der Wiedererkennungs-Variablen leitet über zu der Frage, wie das Zustandekommen des Frequency-Validity-Effekts erklärt wurde. Zwei Arbeiten – Hasher et al. (1977) und Bacon (1979) – bemühten sich um eine theoretische Aufarbeitung des Effekts.

Erklärungsansätze

Die konzeptuellen Vorstellungen von Hasher et al. lassen sich in wenigen Sätzen zusammenfassen. Ausgangspunkt ist die Annahme, daß Menschen äußerst kompetent und zuverlässig Informationen über Ereignishäufigkeiten verarbeiten, speichern und abrufen können. Die Güte dieser frequentistischen Information rührt daher, daß sie „automatisch“ enkodiert wird. Besondere Bedeutung kommt dem Wissen über Ereignishäufigkeiten aus zwei Gründen zu. Erstens ermöglichen Informationen über Ereignishäufigkeiten, alte und neue Ereignisse zutreffend voneinander zu unterscheiden. Zweitens werden Ereignishäufigkeiten als „criterion of certitude“ (Hasher et al., 1977, S. 112) in Situationen genutzt, in denen keine hinreichende Information zur Verfügung steht, um den Wahrheitswert von Aussagen eindeutig zu beurteilen. Konkret bedeutet dies, daß man bei der wiederholten Darbietung von Aussagen, die zwar plausibel erscheinen, deren Wahrheitswert man aber nicht eindeutig beurteilen kann, auf Wissen über die Darbietungshäufigkeit der Items rekurriert, um daraus ein Konfidenz-Urteil abzuleiten. Es gilt: Je größer die erinnerte Darbietungshäufigkeit einer Aussage, desto größer die subjektive Überzeugung in deren Richtigkeit. Ob die Beziehung zwischen Häufigkeit („frequency“) und subjektiver Sicherheit („validity“) linear oder logarithmisch, negativ oder positiv beschleunigt gedacht werden muß, wird von Hasher et al. (1977) nicht diskutiert.

Die zentrale Annahme dieser theoretischen Überlegungen, Ereignishäufigkeiten werden automatisch und damit unter minimalem Verbrauch der Ressource Aufmerksamkeit enkodiert⁴, wurde durch die Formulierung von sechs Bedingungen operationalisiert, die gemeinsam erfüllt sein müssen (Hasher & Zacks 1979, 1984). Eine der sechs Bedingungen, angewendet auf Häufigkeitsinformationen, lautet: Menschen verfügen über Wissen von Ereignishäufigkeiten, das in seiner Qualität unabhängig davon ist, ob die Aufmerksamkeit anläßlich der Enkodierung bewußt („intentional“) oder nur beiläufig auf die Häufigkeit der Ereignisse gerichtet war. Die Frage, inwieweit diese und die anderen Bedingungen empirisch zutreffend sind, löste eine Vielzahl von Arbeiten aus. Obgleich Hasher, Zacks und andere (z.B. Hasher

⁴ „Automatisch“ ist Teil eines Begriffspaars, das von Posner und Snyder (1975) sowie Schneider und Shiffrin (1977) geprägt wurde. Diese Autoren unterscheiden mentale Vorgänge danach, ob sie die begrenzt verfügbare Aufmerksamkeit nur minimal („automatische“ Prozesse) oder wesentlich („kontrollierte“ Prozesse) in Anspruch nehmen.

& Chromiak, 1977; Hasher & Zacks, 1979; Zacks, Hasher & Sanft, 1982; Hasher & Zacks, 1984) eine Fülle empirischer Evidenz zugunsten der automatischen Enkodierungs-Hypothese von Ereignishäufigkeiten zusammengetragen konnten, ist der Status dieser Hypothese gegenwärtig noch umstritten. So wurde z.B. von anderen Autoren bezweifelt, ob es für die Güte von Häufigkeitsinformation tatsächlich irrelevant ist, ob sie intentional oder nur beiläufig enkodiert wurde (z. B. Rose & Rowe, 1976; Greene, 1984, 1986; Naveh-Benjamin & Jonides, 1986; Maki & Ostby, 1987).

Bacon (1979) teilt nun die Ansicht von Hasher et al. (1977), insofern auch diese Autorengruppe ein im Gedächtnis repräsentiertes „*criterion of certitude*“ annimmt. Dieses Kriterium resultiere aber nicht aus einer Schätzung der Darbietungshäufigkeit der Aussagen, aus der dann ein Urteil über die Wahrheit der Aussage abgeleitet wird. Das entscheidende Kriterium wird direkt durch ein Wiedererkennungs-Urteil, in dem eine Aussage entweder als wiederholt („alt“) oder als nicht-wiederholt („neu“) klassifiziert wird, zur Verfügung gestellt. Bacon (1979, S. 242) glaubt: „... *that critical function is served directly by the subject's memory judgment concerning the test statement being rated, without the mediated and redundant process of estimating the frequency with which the statement has occurred.*“

Ohne diesen Aspekt hier zu vertiefen, sei der Hinweis gestattet, daß die Basonsche Annahme, wonach nicht die Schätzung über die Darbietungshäufigkeit sondern das subjektive „alt“/„neu“-Urteil die materielle Grundlage der Frequency-Validity-Beziehung darstelle, Fragen danach aufwirft, welche Gedächtnisrepräsentationen dem Häufigkeits- und dem „alt“/„neu“-Urteil zugrunde liegen und in welcher Beziehung diese Urteile zueinander stehen. Diese Fragen werden in der Literatur kontrovers diskutiert. Sowohl der Standpunkt, beide Urteile stammen aus der gleichen Gedächtnisrepräsentation, wie auch der Standpunkt, beide Urteile stammen aus getrennten Gedächtnisrepräsentationen, werden vertreten (siehe hierzu z.B. Underwood, 1969; Underwood, Zimmerman & Freund, 1971; Hintzman, 1976, 1988; Hintzman & Block, 1971; Harris, Begg & Mitterer, 1980; Begg, Maxwell, Mitterer & Harris, 1986).

Soweit nun das Streiflicht über die Vorstellungen zum Frequency-Validity-Effekt. Allerdings soll dieser Kurzüberblick nicht enden, ohne vorher noch den in der sozialpsychologischen Literatur diskutierten Exposure-Effekt zu erwähnen, der mit dem Frequency-Validity-Effekt zumindest phänomenologisch eng verwandt ist. Konfrontiert man Menschen wiederholt mit Stimuli, die für sie ursprünglich neu waren, so entwickeln sie im Laufe der Darbietungen zunehmend positive Affekte gegenüber diesen Stimuli. Einen hervorragenden Überblick über diesen Forschungsbereich bietet die Arbeit von Bornstein (1989). Beachtenswert ist die Forschung zum Exposure-Effekt deshalb, weil hier gleichfalls die Frage nach der Bedeutung der subjektiven Wiedererkennung eines wiederholten Stimulus im Mittelpunkt einer Kontroverse stand, die zwischen Moreland und Zajonc (1977, 1979; Zajonc, 1981) auf der einen Seite und Birnbaum und Mellers (1979a,b; Birnbaum, 1981) auf der

anderen Seite ausgetragen wurde. Der Kern der Diskussion bestand in der Frage, ob die positiven affektiven Konsequenzen (z.B. erhoben in Form von „liking“ Ratings) der Wiederholung eines Stimulus einer *kognitiven* Moderation bedürfen, die in der Wiedererkennung des Stimulus zum Ausdruck käme (Position von Birnbaum), oder ob die affektiven Konsequenzen eine direkte Funktion der bloßen Darbietung des Stimulus seien und Wiedererkennung somit keine notwendige Bedingung für die Ausbildung des Exposure-Effekts wäre. Es erstaunt sicherlich nicht zu hören, daß die empirische Basis zu dieser Kontroverse gleichfalls verworren und widersprüchlich ist (nachzulesen z.B. in: Russell & Woudzia, 1986; Brooks III & Watkins, 1989, siehe aber auch Bornstein's (1989) Resümee zu Arbeiten, die unterschwellige Reize verwendeten).

Hindsight-Bias:

Methoden, Ergebnisse und Erklärungsansätze

Methoden und Ergebnisse

Da der derzeitige Forschungsstand zum Hindsight-Bias bereits in verschiedenen Überblicksartikeln dokumentiert ist (siehe z.B. Hawkins & Hastie, 1990; Christensen-Szalanski & Willham, 1991), möchte ich mich auf wenige Stichworte zum Untersuchungsmaterial, zu den Methoden und einigen ausgewählten Ergebnissen beschränken. Verzerrungen der Erinnerung im Sinne des Hindsight-Bias konnten bei völlig unterschiedlichen Ereignissen, die es zu beurteilen galt, demonstriert werden. Sowohl bei geschichtlichen Ereignissen (Fischhoff, 1975, 1980), medizinischen Diagnosen (z.B. Arkes, Faust, Guilmette & Hart, 1988), wissenschaftlichen Experimenten (Davies, 1987), wie auch bei Almanachfragen (z.B. Campbell & Tesser, 1983; Fischhoff, 1977) manifestierte sich der Effekt: Im Mittel bewegten sich die Personen in ihren Erinnerungs-Urteilen auf die erhaltene Information zu. In den Kontroll-Aussagen, in denen keine Information gegeben wurde, fand diese systematische Verzerrung der Erinnerungs-Urteile nicht oder nur in viel geringerem Maße statt. Der Bias (d.h. die Effektgröße) war bei der Beurteilung von Almanachfragen größer als bei der Prognose des Ausgangs bestimmter Ereignisse, wie z.B. historischer Schlachten oder medizinischer Diagnosen (Christensen-Szalanski & Willham, 1991). Unterschiede in der Größe des Effekts traten auch in Abhängigkeit davon auf, welche experimentelle Anordnung verwendet wurde. Der Hindsight-Bias-Effekt ist bei einer „Gedächtnis“-Instruktion in der Regel kleiner als bei einer „hypothetischen“ Instruktion (siehe Fischhoff, 1977). Bei einer Gedächtnis-Instruktion werden die Versuchspersonen zunächst gebeten, ein Ereignis zu beurteilen, bevor sie in der zweiten Phase Informationen über den vorab beurteilten Gegenstand erhalten. In der dritten Phase werden sie dann instruiert, sich an ihr ursprüngliches Urteil zu erinnern. Die

Vorgehensweise bei einer hypothetischen Instruktion besteht darin, den Personen zunächst Informationen über das Ereignis zur Verfügung zu stellen. Anschließend werden sie gebeten, sich vorzustellen, wie sie das Ereignis ohne Kenntnis der Lösung beurteilt hätten. Die „hypothetischen“ Urteile werden dann mit den Urteilen einer Kontrollgruppe, die keine Information über das Ereignis erhielt, verglichen. Ein Grund für die unterschiedlichen Effektgrößen besteht sicherlich darin, daß die hypothetische im Vergleich zur Gedächtnisinstruktion andere kognitive Leistungen erfordert. Bedarf es im einen Fall einer Erinnerungsleistung an ein vorab gefälltes Urteil, muß im anderen Fall vorhandenes Wissen außer Acht gelassen werden, um zu einem hypothetischen Urteil gelangen zu können.

In mehreren Studien zum Hindsight-Bias wurde, beabsichtigt oder unbeabsichtigt, die Stärke der Gedächtnisspuren, die über das ursprüngliche Urteil oder die Information gebildet wurden, manipuliert. Mußten die Versuchspersonen z.B. ihr ursprüngliches Urteil begründen (Hell, Gigerenzer, Gauggel, Mall & Müller, 1988) oder mußten sie eigene Antwortalternativen generieren (Pennington, 1981), verringerte dies den Hindsight-Bias. Erklären kann man diese Ergebnisse durch eine vermutlich tiefere Enkodierung des ursprünglichen Urteils. Analog konnte Wood (1978) durch mehrmalige Darbietung der Information einen größeren Hindsight-Bias erzielen: Die Wiederholung ging vermutlich mit einer tieferen Enkodierung einher.

Erklärungsansätze

Nach Hawkins und Hastie (1990) wurden bislang vier allgemeine Strategien – *direkte Erinnerung an das ursprüngliche Urteil, anchoring and adjustment-Prozesse, Neubeurteilung* und *motivational bedingte Adjustierungsprozesse* – identifiziert, die dem Erinnerungs-Urteil innerhalb eines Hindsight-Bias-Designs zugrundeliegen könnten.

Die erste Strategie besteht darin, das ursprüngliche Urteil direkt aus dem Langzeitspeicher abzurufen. Nach Schätzungen von Fischhoff (1977) und Wood (1978) gelingt dies in ungefähr zwei Drittel aller Urteile. Allerdings dürften diese Schätzungen mit Fehlern behaftet sein, da aus der Identität von ursprünglichem Urteil und Erinnerungs-Urteil nicht zwingend geschlossen werden kann, daß das letzte Urteil direkt aus dem Langzeitspeicher abgerufen werden konnte. Da der Hindsight-Bias nicht aufträte, sobald alle Urteile direkt erinnert werden könnten, liegt die Vermutung nahe, daß das Vergessen des ursprünglichen Urteils eine notwendige Bedingung zur Ausbildung des Hindsight-Bias ist (Hell et al., 1988).

Der zweiten Strategie zu einem Erinnerungs-Urteil zu gelangen, liegen anchoring and adjustment-Prozesse zugrunde. Zunächst wird die nach dem Erhalt der Information geformte Wissensbasis und subjektive Sicherheit als Ausgangspunkt gewählt. Dieser Ausgangspunkt bzw. Anker wird dann sukzessive – z.B. durch die Überlegung „Was wußte ich denn bereits damals?“ – angepaßt, um zu einem Erinnerungs-Urteil zu gelangen, welches das

ursprüngliche Wissen und die ursprüngliche subjektive Sicherheit über das Ereignis widerspiegelt (Fischhoff, 1975; Fischhoff & Beyth, 1975). Wird dieser Adjustierungsprozeß nur ungenügend durchgeführt, manifestiert sich in Folge der Hindsight-Bias. Verschiedene Aspekte sprechen gegen das Konzept von anchoring and adjustment-Prozessen. So vertreten zum einen Hawkins und Hastie (1990) sowie Fischhoff (1977, S. 357) die Ansicht, daß diese Strategie für wahre und falsche Ereignisse zu gleich großen Effekten führen sollte. Tatsächlich ist der Hindsight-Bias für wahre Ereignisse aber größer (siehe z.B. Fischhoff, 1975, 1977; aber auch Hasher, Attig & Alba, 1981). Und zum anderen verbirgt das anchoring and adjustment-Konzept mehr als es zu entschlüsseln scheint, denn wie der Adjustierungsprozeß genau vonstatten gehen soll und warum dieser Prozeß systematische Verzerrungen mit sich bringt, wird nicht eindeutig konzeptualisiert.

Eine dritte Strategie, das ursprüngliche Urteil zu rekonstruieren, besteht in der Neubeurteilung des Ereignisses. Hawkins und Hastie (1990) untergliedern diesen erneuten Urteilsprozeß in drei Stufen: (1) Informations-Sammlung, (2) Informations-Bewertung und (3) Informations-Integration.

- (1) Informations-Sammlung: Die Verzerrung der Erinnerungs-Urteile im Sinne des Hindsight-Bias kommt auf dieser Stufe des Urteilsprozesses dadurch zustande, daß nach dem Erhalt der Information die Wissensbasis, die dem ursprünglichen Urteil zugrundelag, nur noch selektiv erinnert wird. Der Zugang zu den Teilen der Wissensbasis, die inkompatibel mit der erhaltenen Information sind, ist verbaut oder zumindest erschwert.
- (2) Informations-Bewertung und (3) Informations-Integration: Bei diesen beiden Stufen des erneuten Urteilsprozesses handelt es sich nach Hawkins und Hastie (1990) um zwei konzeptuell eng verwandte Mechanismen.

Im Fall der Informations-Bewertung erfährt nach dem Erhalt der Information nur das unmittelbar relevante und aufgabenspezifische Wissen eine „Überarbeitung“, so daß ein kohärentes Modell über das Ereignis entsteht. Diese Überarbeitung besteht z.B. darin, die erhaltene Information in die Wissensbasis einzubauen, indem kausale Beziehungen zwischen schon vorher vorhandenem Wissen und der Information konstruiert werden. Auch im Mechanismus der Informations-Integration findet eine Überarbeitung des vorhandenen Wissens statt, die allerdings umfassender ist. Jetzt wird nicht nur das mentale Modell über das zu beurteilende Ereignis (z.B. der Krieg zwischen den Briten und den nepalesischen Gurkas), sondern das mentale Modell über eine ganze Klasse von Ereignissen (z.B. Kolonialkriege) revidiert. Dies könnte z.B. bedeuten, daß nach der Information *Die Briten haben diese koloniale Auseinandersetzung gewonnen* Hinweisreize wie *moderne Waffentechnik, Truppenstärke* etc. höher gewichtet werden, wohingegen andere Hinweisreize wie z.B. *Ortskenntnis* an Aussagekraft verlieren.

Die letzte Strategie ist durch den Einfluß motivationaler Variablen geprägt. So könnte z.B. das Bedürfnis nach vorteilhafter Selbstdarstellung oder aber eine hohe Ich-Involvierung im Hinblick auf das Aufgabenmaterial ein im Lichte der Information besseres Urteil initiieren als ursprünglich getroffen

wurde. Allerdings ist, wie Hawkins und Hastie anmerken, auch die Erklärungskraft des motivationalen Ansatzes eingeschränkt, da die Effekte der bislang untersuchten motivationalen Variablen eher gering sind oder nur in Interaktion mit kognitiven Variablen zum Tragen kommen (Hell et al., 1988) und eine konzeptuelle Trennung zwischen motivationalen und kognitiven Variablen äußerst schwierig ist.

Soweit der an Hawkins und Hastie (1990) angelehnte Überblick über die theoretische Diskussion zum Hindsight-Bias. Aus der vorgeschlagenen Systematisierung werden die Prozesse *direkte Erinnerung an das ursprüngliche Urteil* und *Neubeurteilung* nachfolgend wieder aufgegriffen, wenn auch in ein anderes theoretisches Vokabular gekleidet.

Integratives Modell für den Frequency-Validity-Effekt und den Hindsight-Bias: Vorbemerkungen

Explanandum

Bei beiden Effekten war der Erklärungsgegenstand bislang eine Differenz. Im Falle des Frequency-Validity-Effekts besteht diese in der durchschnittlichen, d.h. über Individuen hinweg gemittelten Differenz zwischen sukzessiven Konfidenz-Urteilen. Im Hindsight-Bias konzentrierten sich die Erklärungsbemühungen auf die durchschnittliche Differenz zwischen Konfidenz-Urteil und Erinnerung. Unbeachtet blieben in beiden Fällen die kognitiven Prozesse, die zum Konfidenz-Urteil in der ersten Sitzung führen. Das erste Konfidenz-Urteil diente lediglich als Ausgangswert, mit dessen Hilfe die Größe der jeweilig interessierenden Differenz bestimmbar wurde.

Im vorliegenden Modell werden nun nicht nur die Differenzen, sondern auch die Vorstufe der Differenzen – die Urteile selbst – in das Explanandum einbezogen. Zwangsläufig richtet sich dann der Blick auf die kognitiven Prozesse, die diesen Urteilen zugrundeliegen. Berücksichtigt man weiterhin, daß infolge der experimentellen Logik in sukzessiven Sitzungen verschiedene experimentelle Randbedingungen realisiert sind, liegen den Urteilen der einzelnen Sitzungen zumindest zum Teil unterschiedliche kognitive Prozesse zugrunde. Natürlich mag es zunächst spitzfindig erscheinen, zwei getrennte Urteile anstelle der Differenz der Urteile als Explanandum zu betrachten. Gleichwohl aber bietet diese Vorgehensweise Vorteile. Erstens erlaubt sie, zwei phänomenologisch unterschiedliche Effekte – Frequency-Validity und Hindsight-Bias – innerhalb eines gemeinsamen Modells (und einer gemeinsamen Sprache) zu diskutieren, weil die Urteile, so die zentrale Annahme, Ausdruck prinzipiell gleicher Konstruktions-, Rekonstruktions- und Erinnerungsprozesse sind. Zweitens können diverse Bedingungsvariablen, die die Effekte moderieren, in das Modell integriert werden, weil die Urteile nicht

mehr nur als Konsequenz des Wirkens einer einzigen Variablen, sondern mannigfacher kognitiver Prozesse verstanden werden. Die theoretische Diskussion endet dann nicht mehr mit dem Streit – Hasher oder Bacon – sondern geht über in die Frage: Welcher Prozeß fand wie häufig statt? Drittens schärft dieses Modell den Blick dafür, daß den Urteilen einer Person unterschiedliche kognitive Prozesse zugrundeliegen, wie auch dieselbe Aufgabe von mehreren Personen gänzlich unterschiedlich gelöst werden kann. Die Aggregation der Urteile über Personen und über Aufgaben birgt die Gefahr, daß die Variabilität der kognitiven Prozesse aus dem Blick gerät.

Grenzen

Nach diesen – noch einzulösenden – Versprechungen aber auch gleich die Frage: Was kann das Modell nicht leisten, wo sind seine Grenzen? Es kann nur bedingt den *unterschiedlichen Antwortformaten*, der *Verschiedenartigkeit des Aufgabenmaterials* und den *verschiedenen Instruktionen im Hindsight-Bias-Design* Rechnung tragen.

In einer Frequency-Validity-Untersuchung besteht das *Antwortformat* immer in einem Urteil über die subjektive Sicherheit, also einem Konfidenz-Urteil. In einer Untersuchung zum Hindsight-Bias kann das Antwortformat entweder in einem Konfidenz-Urteil oder in der Schätzung einer physikalischen Größe (z.B. Metermaß: Höhe des Eiffelturms) bestehen. Um eine gemeinsame Betrachtung von Frequency-Validity-Effekt und Hindsight-Bias zu ermöglichen, war es notwendig, eine beiden Effekten gemeinsame abhängige Variable zu wählen. Aus diesem Grund werden in diesem Modell nur Konfidenz-Urteile diskutiert, Urteile über physikalische Größen sind ausgeklammert.

Bemüht man sich um eine Klassifikation des *Aufgabenmaterials*, das in den verschiedenen Untersuchungen verwendet wurde, dann eröffnet sich eine Ordnungsdimension durch die Frage, welcher Art denn die Aussage sei, die von den Personen beurteilt werden solle. Beinhaltet die Aussage eine quantitative Variable, wie z.B. eine Einwohnerzahl in *München hat mehr als 800.000 Einwohner*, könnte man sie als quantitative (komparative) Aufgabe kennzeichnen. Beinhalten Aussagen hingegen einen qualitativen Sachverhalt, wie z.B. eine Krankheit in *Dostojewski litt an Epilepsie*, lassen sie sich als qualitative Aufgaben typisieren. Innerhalb der quantitativen Aufgaben kann eine weitere Unterscheidung getroffen werden. In der Aussage *München hat mehr Einwohner als Dortmund* sind zwei Elemente (München und Dortmund) auf einer Dimension (Einwohnerzahl) zu beurteilen. In *München hat mehr als 800.000 Einwohner* ist lediglich die Beurteilung eines Elements auf einer Dimension erforderlich. Die quantitativen Aufgaben lassen sich also noch danach unterscheiden, ob sie ein oder zwei Elemente umfassen. Auf der Grundlage dieser Unterscheidung ist jetzt eine Aussage darüber möglich, für welche Art von Aufgaben die theoretischen Überlegungen des Modells Gültigkeit besitzen sollen. Daher hier gleich die zweite Einschränkung. Qualitative Aufgaben werden im Folgenden völlig ausgeklammert, da es mir nicht gelang,

empirisch gehaltvolle und testbare Aussagen über das Urteilsverhalten bei diesem Aufgabentyp zu machen.

Die dritte Einschränkung bezieht sich auf die Unterscheidung zwischen *hypothetischer* und *Gedächtnisinstruktion* im Hindsight-Bias-Design. Dieses Modell bemüht sich lediglich um die Beschreibung der Prozesse, die einer Gedächtnisinstruktion zugrunde liegen könnten.

Noch eine letzte Bemerkung zu den Grenzen. Der Status des Modells ist in vielen Aspekten heuristischer Art, d.h. die Übersetzung der Überlegungen in ein testbares Modell gilt es noch zu leisten. Die Frage, nach dem „wie“ der Übersetzung und den damit einhergehenden Problemen wird ausführlicher im letzten Abschnitt diskutiert.

Theoretische Anleihen

Zur Modellierung des Urteilsverhaltens in der ersten Sitzung knüpft das vorliegende Modell an die Theorie der „Probabilistischen Mentalen Modelle“ (PMM) an (Gigerenzer, Hoffrage & Kleinböling, 1991). Diese Theorie, die in der Tradition der „Overconfidence“-Forschung steht (siehe Sammelreferate hierzu: Lichtenstein, Fischhoff & Phillips, 1982; O'Connor, 1989), ist die bislang einzige Theorie, in der versucht wird, die kognitiven Prozesse, die einem Konfidenz-Urteil zugrundeliegen, testbar zu modellieren. Formuliert wurde die PMM-Theorie anhand von Fragen mit zwei Antwortalternativen, z. B. *Welche Stadt hat mehr Einwohner? (a) Stuttgart (b) Mannheim*. Bei Aufgaben dieser Art beurteilen die Versuchspersonen – nachdem sie sich für eine Alternative entschieden haben – auf einer Konfidenz-Skala, die von 50% bis 100% (Abstufung in 10%-Schritten) reicht, ihre subjektive Sicherheit bezüglich der Richtigkeit ihrer Entscheidung. Die PMM-Theorie läßt sich aber auch auf quantitative Aussagen, in denen ein Vergleich von zwei Elementen auf einer Dimension stattfindet, z.B. *Stuttgart hat mehr Einwohner als Mannheim*, ausdehnen. Bei quantitativen Aussagen dieser Art beurteilen die Versuchspersonen ihre subjektive Sicherheit auf einer Konfidenz-Skala, die von 0% bis 100% reicht (Abstufung in 10%-Schritten). In den Erläuterungen zum vorliegenden Modell werde ich mich auf quantitative Aussagen mit zwei Elementen und die zugehörige Konfidenz-Skala (0%-100%) beziehen.

Der folgende Abschnitt enthält eine Skizze der kognitiven Prozesse, die dem Konfidenz-Urteil der ersten Sitzung einer Frequency-Validity- oder einer Untersuchung zum Hindsight-Bias zugrundeliegen.

Kognitive Prozesse der ersten Sitzung (1. Konfidenz-Urteil)

Kognitive Prozesse bei 0% und 100%-Urteilen

0% bzw. 100%-Urteile werden dann getroffen, wenn es der Versuchsperson möglich ist, ein sogenanntes „Lokales Mentales Modell“ (LMM) zu generieren. Dies gelingt, wenn sie über sicheres und direktes Wissen zu dem in der Aussage angesprochenen Gegenstand verfügt: Eine geborene Stuttgarterin, die in Mannheim studierte, hat möglicherweise direkte Kenntnisse über die Einwohnerzahlen beider Städte, so daß sie eine eindeutige Entscheidung treffen kann. 0% oder 100%-Urteile erfolgen aber auch, wenn auf der Grundlage deduktiver Inferenzen Eindeutigkeit hergestellt werden kann.

Kognitive Prozesse bei 10%-40% und 60%-90% Urteilen

Kann kein LMM generiert werden, schließt sich der Versuch an, ein „Probabilistisches Mentales Modell“ zu konstruieren. In diesem Konstruktionsvorgang werden nicht nur die Elemente „Stuttgart“, „Mannheim“ und die Zielvariable „Einwohnerzahl“ berücksichtigt, sondern es wird auch indirektes Wissen nutzbar gemacht, das mit der Aussage *Stuttgart hat mehr Einwohner als Mannheim* assoziiert ist. Dieses indirekte Wissen besteht – in der Begrifflichkeit der PMM-Theorie – in einer Referenzklasse und Wissen über diese Referenzklasse. In dem exemplarischen Städtevergleich *Stuttgart-Mannheim* bestünde die Referenzklasse aus allen größeren deutschen Städten. Über diese Städte besitzen die Versuchspersonen bisweilen fragmentarisches, manchmal umfangreiches Wissen, und gelegentlich ist ihnen nur der Name der Stadt vertraut. Dieses Wissensreservoir kann man sich auch in Form einer Matrix vorstellen. Abgetragen sind in dieser Matrix erstens die Elemente der Referenzklasse, d.h. die Städte und zweitens das Wissen über die Städte in Form von Variablen mit ihren jeweiligen Ausprägungen. Variablen in dem Städtevergleich könnten z.B. *ist Universitätsstadt*, *ist Landeshauptstadt*, *hat einen IC-Bahnhof* sein. Die Aufgabenstellung definiert nun, welche der Variablen die gesuchte Zielgröße ist. Glaubt die Versuchsperson, daß zumindest einige Variablen mit der Zielgröße kovariieren, können diese Variablen als Hilfsvariablen oder *Cues* herangezogen werden. Konkret: Die *Landeshauptstadt*-Variable kann als Cue verwendet werden, wenn die Versuchsperson annimmt, daß Städte, die den Status einer Landeshauptstadt haben, gewöhnlich größer sind als solche, denen dieser Status fehlt. Trifft eine Versuchsperson ein Urteil im Wissen darum, daß Stuttgart, aber nicht Mannheim eine Landeshauptstadt ist, dann hängt die Trefferwahrscheinlichkeit dieses Urteils von der Aussagekraft (=Validität) des herangezogenen Cues für den Städtevergleich ab.

Eine Möglichkeit, wie die Validität eines Cues im konstruierten „Probabilistischen Mentalen Modell“ repräsentiert sein könnte, ist die einer bedingten Häufigkeit. Stellt man mit den Elementen der Referenzklasse alle möglichen

Zweierkombinationen her und vergleicht die Ausprägung des Cues mit der Zielvariablen, dann läßt sich ermitteln, in wieviel Prozent der Fälle der Algorithmus *Wähle die Stadt, welche nach dem Landeshauptstadt-Cue die größere ist* zu richtigen Entscheidungen führt. Differenziert ein Cue nicht, weil beide oder keine der Städte Landeshauptstädte sind, wird diese Paarung bei der „Berechnung“ der Validität eines Cues nicht berücksichtigt.

Konfidenz-Urteile kommen nun nach Gigerenzer et al. (1991) dadurch zustande, daß die urteilende Person auf der Grundlage des aktivierten Cues zunächst eine Entscheidung zugunsten einer Stadt trifft und als Konfidenz in die Richtigkeit der Entscheidung ihre Validitätsschätzung für den aktivierten Cue mitteilt. Von allen zur Verfügung stehenden Cues wird den Autoren zufolge allerdings immer nur derjenige aktiviert, dem die höchste Cue-Validität zukommt. Sollte dieser nicht anwendbar sein, z.B. weil beide Städte Landeshauptstädte sind, wird der Cue mit der zweithöchsten Cue-Validität aktiviert und auf seine Anwendbarkeit getestet. Ist er anwendbar, d.h. differenziert er zwischen den Städten, dann wird dessen Cue-Validität als Konfidenz-Urteil angegeben.

Abschließend noch eine Bemerkung zu der Frage, wie Menschen zu Schätzungen über die Cue-Validitäten gelangen. Ein Ausgangspunkt für das Erlernen der Cue-Validitäten könnte darin bestehen, daß Menschen die Situationen, in denen sie sicheres Wissen über die Ausprägungen der Cues besitzen, als Grundlage heranziehen, um die Cue-Validitäten zu schätzen. Die so aus der Untermenge aller LMM ermittelten subjektiven Cue-Validitäten werden dann auf Situationen generalisiert, in denen kein sicheres Wissen verfügbar ist. Dieser Generalisierung läge die Annahme zugrunde, daß sich die so geschätzten Cue-Validitäten nicht systematisch von den ökologischen, d.h. den tatsächlichen Validitäten unterscheiden, sondern mit diesen korrespondieren. Gigerenzer et al. (1991, S. 510) nehmen mit Hinweis auf entsprechende positive Evidenz aus der Neo-Brunswikschen Forschung grundsätzlich eine Korrespondenz zwischen den geschätzten und den ökologischen Validitäten an.

Kognitive Prozesse bei 50%-Urteilen

Zu einem 50%-Urteil können verschiedene Prozesse führen. Im Folgenden werden drei Prozesse – *kein Cue ist aktivierbar, mehrere Cues werden aktiviert und Aktivierung des Basisraten-Cues* – unterschieden.

Nach der PMM-Theorie werden 50%-Urteile dann gefällt, wenn auch der Cue mit der niedrigsten Validität keine Differenzierung zwischen den Städten ermöglicht. Sollte also *kein Cue aktivierbar* sein, rät die Versuchsperson und gibt als Konfidenz-Urteil 50% an.

Eine Alternative zu der Annahme der PMM-Theorie, daß immer nur ein Cue herangezogen wird, um eine Entscheidung und ein Konfidenz-Urteil zu fällen, besteht in einem Integrations-Algorithmus, demzufolge mehrere Cues in einer Entscheidung verarbeitet werden. Der Vorteil eines solchen Algorithmus

bestünde darin, daß größere Teile des möglicherweise auch widersprüchlichen Wissens, welches in dem eigens zur Aufgabenlösung konstruierten mentalen Modells zur Verfügung steht, nutzbar gemacht werden können. Offen gelassen, wie die Integration mehrerer anwendbarer Cues genau vonstatten geht, könnte am Ende des Integrationsprozesses eine über alle anwendbaren Cues ermittelte Validität von 50% und ergo ein Konfidenz-Urteil von 50% stehen. Hinter dem 50%-Urteil verbirgt sich dann also nicht bloßes Raten, sondern gleichfalls ein Modellierungsprozeß, der durch die *Aktivierung von mehreren Cues* gekennzeichnet ist, die sich aber in ihrer Aussagekraft egalisieren. Ob ein wie auch immer gearteter Integrations-Algorithmus oder die Annahme der PMM-Theorie den tatsächlichen kognitiven Prozessen am nächsten kommt, müssen zukünftige empirische Untersuchungen klären.

Das 50%-Urteil kann aber auch durch die Anwendung nur eines Cues, des *Basisraten-Cues*, zustandekommen. Findet die Versuchsperson z.B. bei dem Städtevergleich Stuttgart-Düsseldorf keine anwendbaren Cues, dann könnte sie die Referenzklasse „Städte“ verlassen und sich fragen: „Was weiß ich denn darüber, wie solch ein Wissens-Quiz zusammengestellt wird?“ Mit dieser Frage ist dann der erste Schritt in Richtung der Konstruktion eines PMM über das Verhalten des Experimentators gemacht. Eine in dieser Situation plausible Vermutung, die auch in anderem Kontext immer wieder bestätigt wird (z.B. Intelligenz- bzw. Leistungstests), wäre, daß die Basisrate wahrer und falscher Aussagen jeweils 50% betrage. Aus dem Basisraten-Cue folgt dann für den Fall, daß sonst keinerlei Wissen über die Aussage zur Verfügung steht, ein Konfidenz-Urteil von 50%.

Kognitive Prozesse der zweiten Sitzung (2. Konfidenz-Urteil oder Erinnerungs-Urteil)

Kurz zur Erinnerung: In einer Frequency-Validity-Untersuchung sollen die Versuchspersonen nach der ersten Sitzung eine bereits einmal beurteilte Aussage erneut einschätzen, d.h., ein zweites Konfidenz-Urteil wird eingefordert. In einer Untersuchung zum Hindsight-Bias wird eine Erinnerung an das ursprüngliche Konfidenz-Urteil erbeten. Trotz dieser unterschiedlichen Anforderungen an die Versuchspersonen sollen sowohl das Zustandekommen des erneuten Konfidenz-Urteils wie auch des Erinnerungs-Urteils mit den gleichen kognitiven Prozessen erklärt werden. In diesem Abschnitt werden lediglich die einzelnen Prozesse erläutert, im nächsten Abschnitt wird dann das Augenmerk auf das Zusammenwirken der Prozesse gerichtet.

Kognitive Prozesse, die zu einer identischen Reproduktion des ersten Konfidenz-Urteils führen

Der Fall einer identischen Reproduktion (d.h., das zweite Konfidenz-Urteil bzw. das Erinnerungs-Urteil ist gleich dem ersten Konfidenz-Urteil) kann durch zwei Prozesse – *direkte Erinnerung, erneute Konstruktion eines mentalen Modells* – erklärt werden.

In einer Frequency-Validity-Untersuchung kann *direkte Erinnerung* an das Erst-Urteil durch die Bemühung um konsistentes und optimiertes Urteilsverhalten („Ich denke nicht erneut über den Sachverhalt nach, sondern versuche, mich an mein letztmaliges Urteil zu erinnern, da ich zwischenzeitlich keine neuen Informationen erhalten habe.“) motiviert sein. Eine notwendige Bedingung für das Streben nach direkter Erinnerung ist allerdings, daß die Versuchsperson eine Aussage als „alt“, d.h. wiederholt zu erkennen glaubt. In einer Untersuchung zum Hindsight-Bias wird der Prozeß der direkten Erinnerung durch die Instruktion („Wie haben Sie zuerst geurteilt?“) initiiert. Ist ein direkter Abruf des ursprünglichen Urteils möglich, so wird dieses exakt reproduziert – unabhängig davon, ob auch die erhaltene Information über den Wahrheitsgehalt der Aussage direkt erinnert werden kann.

Obgleich angestrebt, kann der Versuch, das Urteil direkt zu erinnern, auch fehlschlagen. Dies wird dann der Fall sein, wenn die Gedächtnisrepräsentation des Erst-Urteils aufgrund natürlicher Vergessensprozesse zerfallen oder nicht mehr zugänglich ist oder wenn aufgrund eines falschen Wiedererkennungsurteils („alt“) ein Konfidenz-Urteil zu erinnern versucht wird, das nie – zumindest im experimentellen Kontext – gefällt wurde. Mißlingt also die direkte Erinnerung, dann bemüht sich die Versuchsperson, die Wissensbasis zu errichten, die dem ersten Konfidenz-Urteil zugrundelag, d.h., sie *konstruiert erneut ein mentales Modell*. Gelingt es ihr bereits in der ersten Sitzung, ein LMM zu bilden, ist es wahrscheinlich, daß das notwendige Wissen oder die notwendigen Deduktionen auch jetzt wieder verfügbar sind, so daß das zweite Konfidenz- und das Erinnerungs-Urteil mit dem ersten Konfidenz-Urteil identisch sind. Gelingt hingegen lediglich die Bildung eines PMM, dann ist es nicht sicher, daß exakt das Erst-Urteil reproduziert werden kann. Diese Unschärfe entsteht dadurch, daß die Arbeitsschritte, die zur Konstruktion eines PMM führen (z.B. Errichtung einer Cue-Hierarchie, Schätzung der Validitäten) vermutlich nicht vollständig reliabel sind. Diese letzten Überlegungen leiten unmittelbar zu der nächsten Kategorie kognitiver Prozesse über.

Kognitive Prozesse, die zu einer nicht-identischen Reproduktion des ersten Konfidenz-Urteils führen

Unterschiede zwischen dem Urteil der ersten und zweiten Sitzung können durch zwei Prozesse – *Berücksichtigung neuer Information* oder *erneute Konstruktion eines mentalen Modells* – erklärt werden.

Was passiert, wenn einer Versuchsperson *direkte und sichere Informationen* über den zu beurteilenden Sachverhalt zur Verfügung stehen, die zum Zeitpunkt des Erst-Urteils noch nicht Bestandteil der Wissensbasis waren? Die Versuchspersonen werden aufgrund dieser direkten und sicheren Informationen in der Lage sein, ein LMM zu erstellen. Sollten die erworbenen Informationen⁵ valide sein, werden für wahre Aussagen 100%-Urteile, für falsche Aussagen 0%-Urteile gefällt. In einer Frequency-Validity-Untersuchung ist die Verfügbarkeit neuer Information ein unerwünschtes Phänomen, da die Auswirkung der bloßen Wiederholung der Aussagen auf das Urteilsverhalten untersucht werden soll. Finden die wiederholten Darbietungen aber nicht innerhalb einer Sitzung, sondern über mehrere Tage oder Wochen statt, muß mit einem beabsichtigten oder auch zufälligen Informationserwerb (Lexika, Gespräche mit „Experten“) gerechnet werden. In einer Hindsight-Bias-Untersuchung ist die Informierung über den Wahrheitsgehalt der Aussagen kein unerwünschtes Phänomen, sondern Bestandteil der experimentellen Logik. Die Versuchspersonen werden nach dem ersten Urteil über den Wahrheitsgehalt eines Teils der Aussagen (=Experimental-Aussagen) informiert. Idealerweise erinnern sie sich dann ihrer ersten Urteile, ohne sich dabei durch die erhaltene Information beeinflussen zu lassen.

Können die Versuchspersonen das Erst-Urteil nicht abrufen, besteht die beste Strategie darin, genau das *mentale Modell erneut zu konstruieren*, das zum Erst-Urteil geführt hat. Gleichwohl kann am Ende dieses Rekonstruktionsvorgangs aber ein mentales Modell stehen, das sich vom Original in zweifacher Hinsicht – in der Anzahl der generierten Cues und/oder in der Bewertung des Vertrauheits-Cues – unterscheidet. Werden neue, bislang *noch nicht berücksichtigte Cues generiert* (z.B. aufgrund längerer Latenzzeit, veränderter Suchstrategie nach dem Erhalt der Information im Hindsight-Bias-Design), dann kann diese erweiterte Wissensbasis Auswirkungen auf das zweite Urteil haben. Bedient man sich im Konfidenz- und Erinnerungs-Urteil nur des Cues mit der höchsten Validität, werden die neuen Cues nur dann relevant, wenn sich unter diesen einer befindet, der den bisherigen „Spitzenreiter“ überflügeln kann. Liegt hingegen ein Mittelungs-Algorithmus (= mehrere Cues werden in einem Urteil verarbeitet) zugrunde, wird jeder neu generierte Cue (ungeachtet seiner Cue-Validität) einen Einfluß auf das zweite Urteil besitzen. Welcher Art ist nun der Einfluß der erweiterten Wissensbasis auf das Erinnerungs- bzw. zweite

⁵ „Information“ bezeichnet in Abgrenzung zu unsicherem Wissen, das in der Zwischenzeit verfügbar wurde, sicheres und direktes Wissen, das die Errichtung eines LMM möglich macht. Im Fall von unsicherem Wissen kommt es zur Errichtung eines PMM.

Konfidenz-Urteil? Unterstellt man, daß die von den Versuchspersonen aktivierten, also auch die neu hinzukommenden Cues grundsätzlich valide⁶ sind, dann sollten die zweiten Konfidenz-Urteile in einer Frequency-Validity-Untersuchung für wahre Aussagen anwachsen (d.h., die Vpn werden sicherer, daß die Aussage richtig ist, z.B. 80% anstelle eines Erst-Urteils von 60%), hingegen die Urteile für falsche Items abnehmen (d.h., die Vpn werden sicherer, daß die Aussage falsch ist, z.B. 40% anstelle eines Erst-Urteils von 60%). Analog gilt für eine Hindsight-Bias-Untersuchung, daß unabhängig von der Rückmeldung über den Wahrheitsgehalt die Erinnerungen für wahre (falsche) Aussagen näher an 100% (0%) liegen.

Sowohl in einem Hindsight-Bias wie auch einem Frequency-Validity-Design kann es bei einem inhaltsunabhängigen Cue, dem *Vertrauthets-Cue*, zu einer Neueinschätzung der Cue-Validität und der Anwendbarkeit des Cues kommen. Diese Neueinschätzung ist von der Rückmeldung über den Ausgang des Ereignisses (in Hindsight-Bias-Design) unabhängig. Um diese Überlegung zu erläutern, wird die Unterscheidung zwischen *präexperimenteller* und *experimenteller Vertrautheit* getroffen. Präexperimentelle Vertrautheit bezeichnet die Vertrautheit, die einer Aussage zukommt, noch bevor sie im experimentellen Kontext dargeboten wurde. Diese ist, sofern die Konzepte⁷, die Bestandteile der Aussage sind, für die Versuchsperson völlig neu sind, Null. Die Vertrautheit gegenüber einer Aussage, die durch die Darbietung der Aussage im experimentellen Kontext erzeugt wird, kann als experimentelle Vertrautheit gekennzeichnet werden. *Quellen-Attribution* umschreibt die Zuordnung der Vertrautheit zu der Quelle, aus der sie herrührt. Die Vertrautheit gegenüber einer Aussage kann aus der Quelle „experimenteller Kontext“ („Ich weiß, daß ich diese Aussage in der letzten Sitzung erstmals gehört habe. Davor war mir diese Aussage völlig unbekannt.“), aus der Quelle „präexperimenteller Kontext“ („Ich erinnere mich, über diese Aussage oder einzelne Konzepte der Aussage bereits früher etwas gehört zu haben.“) oder aus beiden Quellen zusammen stammen. Während der ersten Beurteilung einer Aussage kann die Vertrautheitsempfindung noch eindeutig einer Quelle – dem präexperimentellen Kontext – zugeordnet werden. Bereits in der zweiten Darbietung wird die Quellenzuordnung schwieriger, da die Vertrautheitsempfindung aus

⁶ Ob es gerechtfertigt ist, davon auszugehen, daß die neu hinzukommenden Cues grundsätzlich valide sind, ist eine offene Frage, die empirisch geklärt werden muß. Vorstellbar ist auch, daß infolge eines „Confirmation Bias“ (siehe z.B. Klayman & Ha, 1987) nur noch solche Cues neu generiert werden, die die grundsätzliche Tendenz des ursprünglichen Urteils (größer oder kleiner 50%) bestätigen.

⁷ Die Vertrautheit gegenüber einer Aussage setzt sich aus zwei Aspekten zusammen: (a) Konzeptvertrautheit als die Vertrautheit mit den einzelnen Konzepten einer Aussage (z.B.: Ernest Hemingway, 1. Weltkrieg, Fahrer eines Ambulanzwagens) und (b) Satzvertrautheit als die Vertrautheit mit der spezifischen Zusammensetzung der Konzepte, die die Semantik des Satzes konstituieren (z.B.: Ernest Hemingway war im 1. Weltkrieg Fahrer eines Ambulanzwagens). Konzeptvertrautheit ist eine notwendige, aber keine hinreichende Bedingung für die Satzvertrautheit. Letztere ist aber eine hinreichende Bedingung für die Konzeptvertrautheit.

zwei Quellen – dem präexperimentellen und dem experimentellen Kontext – herrühren kann. Fehler, die jetzt auftreten, werden als *Quellen-Fehlattributionen*⁸ bezeichnet, wobei Fehler sowohl in der Zuordnung zur Quelle wie auch in der Einschätzung des Ausmaßes der Vertrautheit pro Quelle auftreten können. Worin besteht die Bedeutung dieser Quellen-Fehlattributionen für unseren Gegenstand?

Aussagen, die vertraut erscheinen, d.h., die in irgendeiner Form mit bereits im Gedächtnis gespeicherter Information korrespondieren, wird ein Wahrheitsbonus eingeräumt (Begg et. al, 1985).⁹ In diesem Sinn wird die präexperimentelle Vertrautheit zu einem Cue für den Wahrheitsgehalt der Aussage. Der „Wahrheitsbonus“ wird aber natürlich nicht gewährt, sobald die Vertrautheit gegenüber einer Aussage ausschließlich aus der Quelle „experimenteller Kontext“ herrührt. In diesem Fall wird der Vertrautheits-Cue als invalide erachtet: „Ich weiß, daß ich diese Aussage das erste Mal in der letzten experimentellen Sitzung gehört habe. Und da ich ferner weiß, daß die Experimentatoren mich sowohl mit wahren, als auch mit falschen Aussagen konfrontieren, und ich zwischenzeitlich nichts Neues erfahren habe, bleibe ich bei meinem Urteil.“ In der Zuordnung der Vertrautheit zu den Quellen „prä-, bzw. experimenteller Kontext“ kann es aber zu Fehlern kommen, von denen zwei für den Frequency-Validity-Effekt Erklärungsrelevanz besitzen. Der erste Fehler besteht darin, daß die Vertrautheitsempfindung anläßlich der zweiten Darbietung einer Aussage fälschlicherweise (auch) der Quelle präexperimenteller Kontext zugeschrieben wird. Damit wird der Vertrautheits-Cue vermeintlich anwendbar. Der zweite Fehler besteht in der rückwirkenden Überschätzung präexperimenteller Vertrautheit, d.h. das Ausmaß der präexperimentellen Vertrautheit wird aufgrund der Fehlattribution der experimentellen Vertrautheit überschätzt und folglich die Validität des Vertrautheits-Cue überbewertet. Beide Fehler können dazu führen, daß das zweite Konfidenz-Urteil höher als das erste ausfällt. Auch in einem Hindsight-Bias-Design sollte die durch die Quellen-Fehlattributionen bedingte post hoc-Überschätzung der präexperimentellen Vertrautheit dazu führen, daß die Erinnerungen an die Konfidenz-Urteile höher ausfallen, als sie tatsächlich waren, und zwar

⁸ Der Begriff „Quellen-Fehlattribution“ wird im Sinne von Lindsay und Johnson (1989) verwendet, die diesen Effekt in der Forschung über die Psychologie von Zeugenaussagen („eyewitness testimony“) diskutieren.

⁹ Natürlich stellt sich die Frage, warum die Versuchspersonen die gespeicherte Information als wahr annehmen und daher aus dem Ausmaß der Korrespondenz zwischen gespeicherter Information und Aussage auf die Wahrheit der Aussage schließen. Möglicherweise, so könnte man spekulieren, liegt die folgende Überlegung zugrunde: Aufgrund der begrenzten Gedächtniskapazität ist man bemüht, nur wahre Informationen zu speichern. Dies ist ein effizientes Selektionskriterium, weil bei multiplen Alternativen aus einer wahren Information zu einem Sachverhalt auf falsche zu schließen ist, umgekehrt aus einer falschen Information seltener auf eine wahre Information geschlossen werden kann.

unabhängig von der Rückmeldung über den tatsächlichen Wahrheitsgehalt der Aussagen.¹⁰

Prozeßmodell und Vorhersagen

Was haben wir mit dem bisher Gesagten gewonnen? Gewonnen haben wir einen Korpus kognitiver Prozesse, mit dem das Urteilsverhalten in einer Frequency-Validity- und einer Hindsight-Bias-Sitzung beschrieben werden kann. In diesem Abschnitt werden diese Prozesse nicht mehr isoliert diskutiert, sondern in ein Prozeßmodell eingeordnet. Zweitens sollen die Vorhersagen über die Konfidenz- und Erinnerungs-Urteile der zweiten Sitzung im Vergleich zum ersten Konfidenz-Urteil in einer Gesamtschau betrachtet werden.

Frequency-Validity-Design

Was genau passiert, wenn in der zweiten Sitzung eine Aussage erneut eingeschätzt werden soll? In Abbildung 1 ist ein Ablauf skizziert, wie er nach dem oben skizzierten Modell aussehen könnte.

Zuerst stellt sich der Urteilende die Frage: *Handelt es sich um eine wiederholte („alte“) oder eine nicht-wiederholte („neue“) Aussage?* Ist der Ausgang dieser Abfrage ein „alt“-Urteil, muß geklärt werden: *Wurde zwischenzeitlich neue und sichere Information über den Aussagengegenstand erworben?* Wenn dies der Fall ist und die Information auch erinnert wird, kann mit dieser neuen Information ein LMM errichtet werden, dem dann ein 0% oder 100% Urteil folgt. Liegt hingegen keine neue Information vor, schließt sich eine weitere Abfrage an: *Kann man sich des Urteils der ersten Sitzung erinnern?* Ist dies möglich, dann wird genau dieses Urteil reproduziert. Eine Erinnerung an das ursprüngliche Urteil findet bei *salienten* Urteilen – 100%, 0% und 50% Urteile – vermutlich häufiger statt als bei 10%-40% und 60%-90%-Urteilen. Kann das ursprüngliche Urteil nicht direkt erinnert werden, wird schließlich erneut ein mentales Modell konstruiert. Zu diesem Prozeß führt auch die erste Abfrage, sofern der Ausgang ein „neu“-Urteil war. Am Ende dieses erneuten Konstruktionsprozesses können mehrere Ergebnisse stehen: Konnte in der ersten Sitzung bereits ein LMM konstruiert werden, sollte dies aufgrund der Verfügbarkeit der sicheren Information auch jetzt wieder gelingen. Folge: Reproduktion des 0% oder 100%-Urteils aus der ersten Sitzung (a). Kommt es in der zweiten Sitzung zu einer analogen Nachbildung des PMM der ersten Sitzung, dann wird gleichfalls das ursprüng-

¹⁰ Natürlich kann auch eine Fehlattribuierung zugunsten der experimentellen Vertrautheit einer Aussage vorkommen (z.B. wird die präexperimentelle Vertrautheit einer Aussage der Quelle experimenteller Kontext zugeordnet). Dieser Fehler wird aber hier vernachlässigt, weil er zu einer Nichtanwendung des Vertrautheits-Cues führen würde und somit für die hier verfolgte Frage keine Erklärungsrelevanz besitzt.

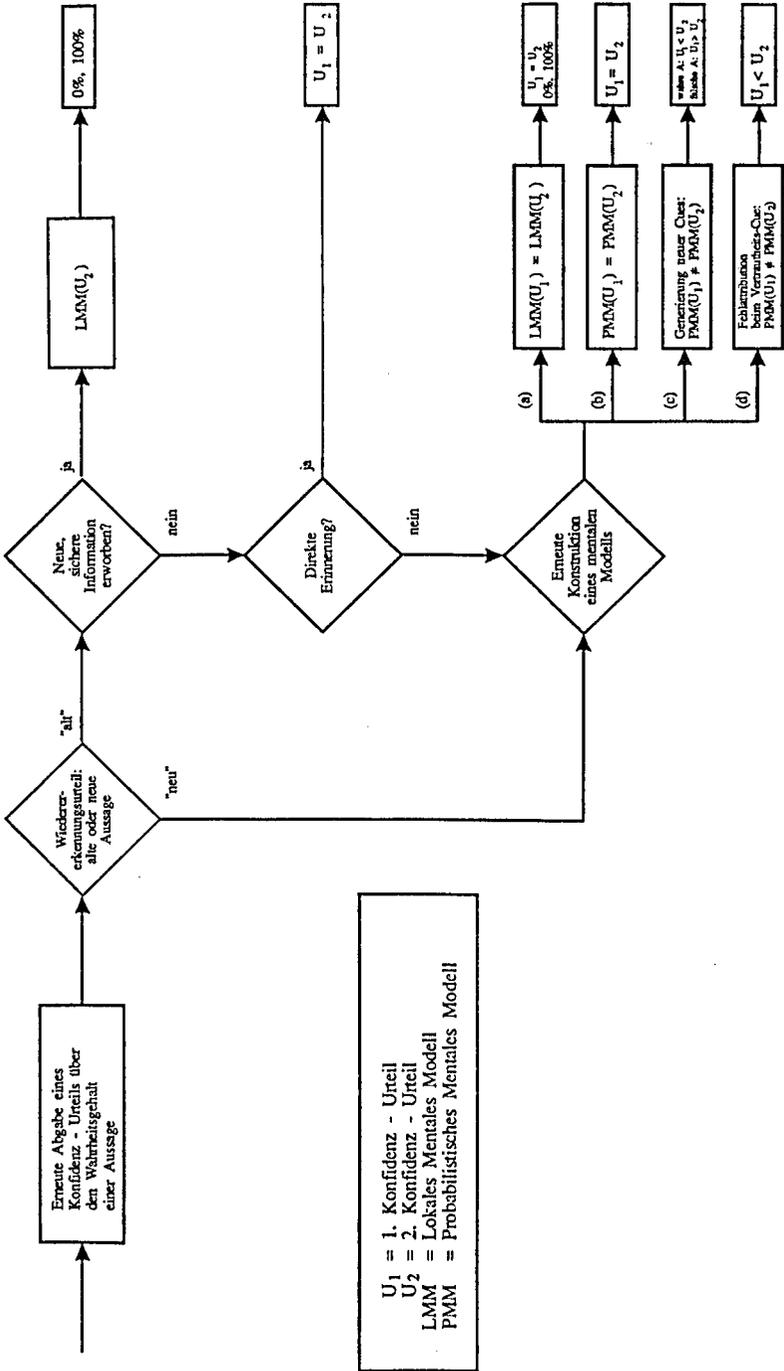


Abbildung 1 Kognitive Prozesse anlässlich des zweiten Konfidenz-Urteils in einem Frequency-Validity-Design.

Tabelle 1 Vorhersagen für das zweite Konfidenz-Urteil in einer Frequency-Validity-Untersuchung in Abhängigkeit der kognitiven Prozesse
 (U_n = Konfidenz-Urteil in der n-ten Sitzung,
 LMM = Lokales Mentales Modell,
 PMM = Probabilistisches Mentales Modell).

1. Sitzung	2. Sitzung	Vorhersagen (im Mittel)	
		wiederholte Aussagen wahr	falsch
Konfidenz-Bereich/ Kognitiver Prozeß	Kognitiver Prozeß (2. Konfidenz-Urteil)		
	0%, 100%-Urteile / Konstruktion eines LMM	- Direkte Erinnerung $U_1 = U_2$	$U_1 = U_2$
	- Erneute Konstruktion eines mentalen Modells: $LMM(U_1) = LMM(U_2)$	$U_1 = U_2$	$U_1 = U_2$
	- Berücksichtigung neuer, sicherer Infomation	$U_2 = 100\%$	$U_2 = 0\%$
10%-40%, 60-90%; 50%-Urteile/ Konstruktion eines PMM; Raten, Integration mehrerer Cues, Basisraten-Cue	- Direkte Erinnerung	$U_1 = U_2$	$U_1 = U_2$
	- Berücksichtigung neuer, sicherer Information	$U_1 = 100\%$	$U_2 = 0\%$
	- Erneute Konstruktion eines mental Modells:		
	1) Analoge Rekonstruktion: $PMM(U_1) = PMM(U_2)$	$U_1 = U_2$	$U_1 = U_2$
	2) Generierung neuer Cues	$U_1 < U_2$	$U_1 > U_2$
	3) Fehlattribution beim Vertrauthets-Cue	$U_1 < U_2$	$U_1 < U_2$

liche Konfidenz-Urteil der ersten Sitzung reproduziert (b). Eine analoge Nachbildung eines PMM wird aber aufgrund der eingeschränkten Reliabilität der diversen Konstruktionsschritte nicht immer gelingen. Die Generierung gänzlich neuer Cues und Fehlattritionen beim Vertrautheits-Cue können dazu führen, daß sich die PMMs der ersten und zweiten Sitzung und folglich die Konfidenz-Urteile der zweiten Sitzung systematisch von den ursprünglichen Urteilen unterscheiden (c, d). Bei der Generierung gänzlich neuer Cues verändert sich das Konfidenz-Urteil der 2. Sitzung in Abhängigkeit vom Wahrheitsgehalt der Aussagen (wahre Aussagen: $U_1 < U_2$; falsche Aussagen: $U_1 > U_2$), wohingegen bei Fehlattritionen des Vertrautheits-Cues unabhängig vom Wahrheitsgehalt der Aussagen das Konfidenz-Urteil ansteigt ($U_1 < U_2$).

In Tabelle 1 werden die Vorhersagen über die Veränderungen der Konfidenz-Urteile von der ersten zur zweiten Sitzung aufgelistet. Um die Tabelle übersichtlich zu gestalten, wurden die Vorhersagen für wahre und falsche Aussagen nicht nach dem Wiedererkennungs-Urteil („alt“, „neu“) unterschieden. Dies auch deshalb, weil diese Unterscheidung nur für einen Prozeß – direkte Erinnerungen – relevant wäre. Die Tabelle liest sich wie folgt: Wurde z.B. in der ersten Sitzung ein PMM konstruiert, d.h. ein 10%-40% oder ein 60%-90% Urteil gefällt (siehe 1. Spalte), dann könnte dieses Urteil in der zweiten Sitzung direkt erinnert werden (siehe 2. Spalte) und folglich entsprächen sich die Konfidenz-Urteile der ersten und zweiten Sitzung (d.h. $U_1 = U_2$, siehe Spalte 3). Allerdings kann es in der 2. Sitzung neben der „direkten Erinnerung“ auch zu anderen Prozessen kommen (z. B. Berücksichtigung neuer Information), die zu entsprechend anderen Vorhersagen im Hinblick auf das Konfidenz-Urteil der zweiten Sitzung führen.

Sowohl der Abbildung 1 wie auch der Tabelle 1 ist zu entnehmen, daß in der zweiten Sitzung einer Frequency-Validity-Untersuchung eine ganze Reihe kognitiver Prozesse stattfinden können, die die Konfidenz-Urteile dieser Sitzung aufklären. Von besonderer Bedeutung für den Frequency-Validity-Effekt ist allerdings der Prozeß „Fehlattrition beim Vertrautheits-Cue“, da es der einzige Prozeß ist, der unabhängig vom Wahrheitsgehalt der Aussagen zu einer Zunahme des Konfidenz-Urteils führt.

Hindsight-Bias-Design

Was passiert, wenn in der zweiten Sitzung das Konfidenz-Urteil der ersten Sitzung erinnert werden soll? Dieser Erinnerungsprozeß wird auf der Grundlage des vorgestellten Modells in Abbildung 2 skizziert.

An erster Stelle steht hier die Frage: *Kann man sich an das Urteil der ersten Sitzung erinnern?* Ist dies der Fall, wird es unmittelbar berichtet. Ähnlich wie oben gilt aber auch hier, daß eine Erinnerung an das ursprüngliche Urteil bei salienten Urteilen (0%, 100%, 50%) vermutlich häufiger stattfindet. Kann das ursprüngliche Urteil hingegen nicht mehr abgerufen werden, dann folgt auch hier die erneute Konstruktion eines mentalen Modells. Motiviert wird dieser Konstruktionsprozeß durch folgende Überlegung: „Vergesse ich für einen

Moment, daß ich nun die Lösung kenne, dann weiß ich jetzt nicht mehr über den zu beurteilenden Gegenstand als zum Zeitpunkt des ursprünglichen Urteils. Folglich beurteile ich diese Aussage mit dem gleichen Wissen und kann hoffen, zu einem identischen Urteil zu gelangen." Folgende Ausgänge können am Ende dieses Konstruktionsprozesses stehen. Konnte in der ersten Sitzung bereits ein LMM konstruiert werden, sollte dies aufgrund der Verfügbarkeit der sicheren Information auch jetzt wieder errichtet werden. Folge: Reproduktion des 0% oder 100%-Urteils aus der ersten Sitzung, das dann als Erinnerungs-Urteil berichtet wird (a). Kommt es in der zweiten Sitzung zu einer analogen Nachbildung des PMM der ersten Sitzung, dann wird gleichfalls das Konfidenz-Urteil der ersten Sitzung reproduziert und als Erinnerungs-Urteil wiedergegeben (b). Die Generierung gänzlich neuer Cues und Fehl-attributionen beim Vertrautheits-Cue können dazu führen, daß sich die PMMs der ersten und zweiten Sitzung unterscheiden und somit die Erinnerungs-Urteile in systematischer Weise von den ursprünglichen Urteilen abweichen (c, d). Natürlich kann es in der erneuten Konstruktion eines mentalen Modells auch vorkommen, daß die Versuchspersonen in der Retrospektive glauben, bereits zum Zeitpunkt des ursprünglichen Urteils über die korrekte Information (die erst später vom Versuchsleiter dargeboten wurde) verfügt zu haben und sich deshalb vermeintlich erinnern, ihr Urteil mit entsprechend hoher Sicherheit gefällt zu haben (e). Diese fehlerhafte Quellenzuordnung der sicheren Information findet aber nur dann statt, wenn keinerlei Erinnerung an das mentale Modell (=PMM) und das Urteil der ersten Sitzung mehr möglich ist.

In Tabelle 2 werden die Vorhersagen über die Veränderungen der Erinnerungs-Urteile im Vergleich zu den ursprünglichen Konfidenz-Urteilen aufgelistet. Die Tabelle ist analog zu Tabelle 1 aufgebaut. Wurde z.B. in der ersten Sitzung ein PMM konstruiert, d.h. ein 10%-40% oder ein 60%-90% Urteil gefällt (siehe 1. Spalte), dann könnte es der Versuchsperson in der zweiten Sitzung gelingen, dieses Urteil direkt zu erinnern (siehe 2. Spalte). Gelingt ihr dies tatsächlich, entspräche das Erinnerungs-Urteil der zweiten Sitzung dem Urteil der ersten Sitzung (d.h. $U_2 = E(U_1)$, siehe Spalte 3). Allerdings kann es in der zweiten Sitzung neben der direkten Erinnerung auch zu anderen Prozessen kommen (z. B. Generierung neuer Cues), die zu entsprechend anderen Vorhersagen im Hinblick auf das Erinnerungs-Urteil der zweiten Sitzung führen.

Was bleibt zu tun?

Soweit das theoretische Gerüst. Aber damit ist nur der erste Teil der Arbeit an diesem Modell getan. Lassen Sie mich diesen Beitrag mit einer Liste der weißen Flecke, die es noch aufzufüllen gilt, beenden.

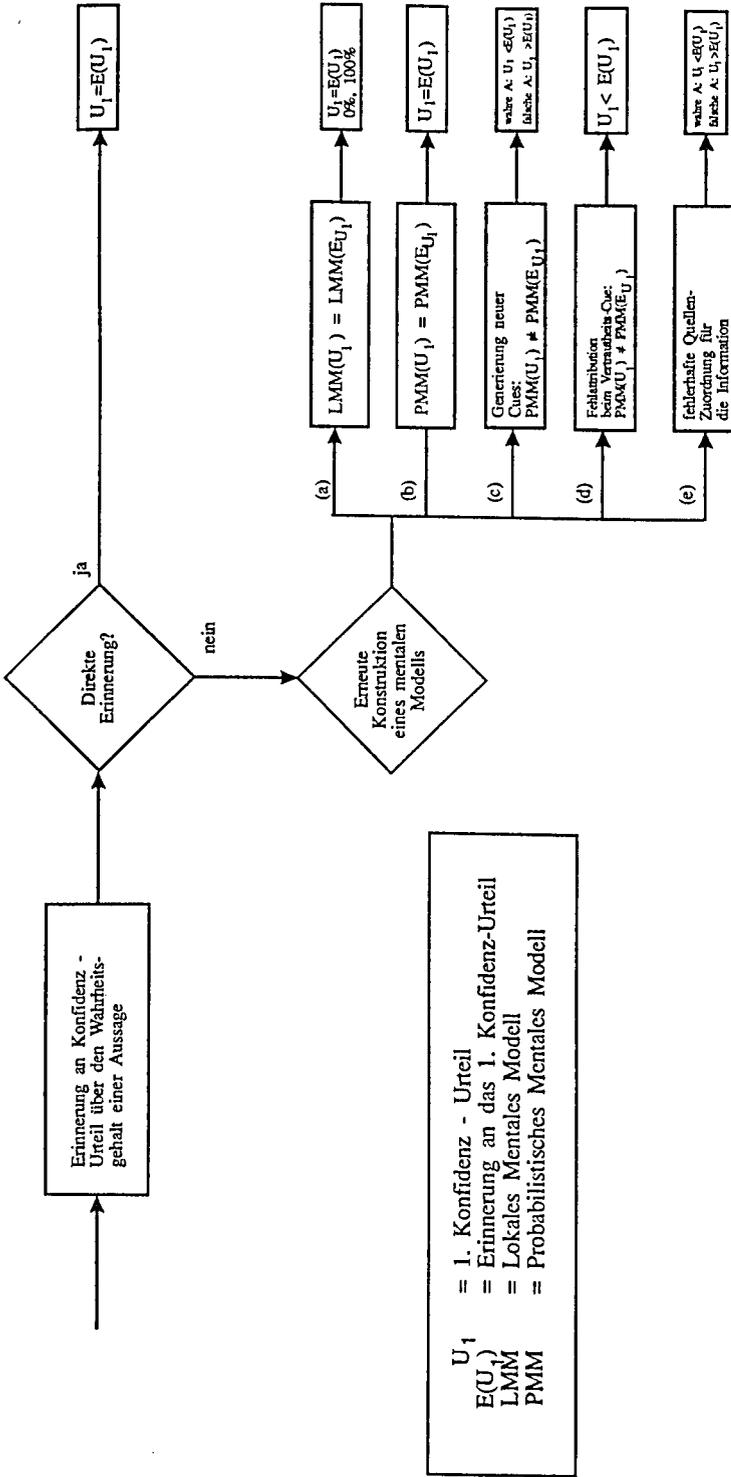


Abbildung 2 Kognitive Prozesse anlässlich des Erinnerungs-Urteils in einem Hindsight-Bias-Design.

Tabelle 2 Vorhersagen für das Erinnerungs-Urteil in einer Hindsight-Bias-Untersuchung in Abhängigkeit der kognitiven Prozesse

(U_1 = Konfidenz-Urteil in der ersten Sitzung,

$E(U_1)$ = Erinnerung an U_1 ,

LMM = Lokales Mentales Modell,

PMM = Probabilistisches Mentales Modell).

1. Sitzung Konfidenz-Bereich/ Kognitiver Prozeß	2. Sitzung Kognitiver Prozeß (2. Konfidenz-Urteil)	Vorhersagen (im Mittel)	
		wiederholte Aussagen wahr	falsch
0%, 100%-Urteile / Konstruktion eines LMM	- Direkte Erinnerung	$U_1 = E(U_1)$	$U = E(U_1)$
	- Erneute Konstruktion eines mentalen Modells: $LMM(U_1) = LMM(U_2)$	$U_1 = E(U_1)$	$U_1 = E(U_1)$
10%-40%, 60-90%; 50%-Urteile/ Konstruktion eines PMM; Raten, Integration mehrerer Cues, Basisraten-Cue	- Direkte Erinnerung	$U_1 = E(U_1)$	$U_1 = E(U_1)$
	- Erneute Konstruktion eines mental Modells:		
	1) Analoge Rekonstruktion: $PMM(U_1) = PMM(U_2)$	$U_1 = E(U_1)$	$U_1 = E(U_1)$
	2) Generierung neuer Cues	$U_1 < E(U_1)$	$U_1 > E(U_1)$
	3) Fehlattribution beim Vertrauheits-Cue	$U_1 < E(U_1)$	$U_1 < E(U_1)$
	4) Fehlerhafte Quellenzuordnung für die Information	$U_1 < E(U_1)$	$U_1 > E(U_1)$

Überprüfung der Modellannahmen

In dieses Modell gehen zwei Gruppen von Annahmen ein, die überprüft werden müssen. Die eine Gruppe umfaßt die Kernannahmen der PMM-Theorie, die gegenwärtig im Kontext der Overconfidence-Forschung am Prüfstein der Empirie gemessen werden (siehe U. Hoffrage in diesem Band). Die Ergebnisse dieser Bemühungen sind für das vorliegende Modell von zen-

traler Bedeutung, da sie den Ausgangspunkt aller nachfolgenden Überlegungen darstellen. Eine zweite Gruppe betrifft die Annahmen, die getroffen wurden, um auf dem theoretischen Hintergrund der PMM-Theorie auch wiederholte Urteile zum gleichen Gegenstand, also Urteile im Rahmen einer Frequency-Validity- und einer Hindsight-Bias-Untersuchung erklären zu können. Beispiele für diese Annahmen sind: (1) Die Anzahl der generierten Cues wird größer, wenn ein Sachverhalt wiederholt eingeschätzt wird. (2) Die erweiterte Wissensbasis führt zu besseren Urteilen über den Wahrheitsgehalt der Aussagen. (3) Fehlattritionen beim Vertrautheits-Cue finden statt und führen gleichfalls zu systematischen Änderungen im Urteilsverhalten.

Effektgröße für jeden Prozeß

Weitere bislang noch unbekannte Größen sind in den Vorhersagen über die Auswirkungen der kognitiven Prozesse enthalten. In den Tabellen 1 und 2 habe ich dafür Ungleichheitszeichen (z.B. $U_1 > U_2$ oder $U_1 < E(U_1)$) eingesetzt. Diese Ungleichheitszeichen vermögen zwar die Richtung der Veränderung, nicht aber die vermutliche Größe des Effekts (d.h. das Ausmaß der Veränderung des Konfidenz-Urteils von der ersten zur zweiten Sitzung) anzudeuten. Wenn die Effektgröße für jeden Prozeß und die Gewichtung für jeden Prozeß (d.h., wie groß ist der prozentuale Anteil der Urteile, der durch den jeweiligen Prozeß zustandekam; wie häufig fand z.B. der Prozeß „Generierung neuer Cues“ in der zweiten Sitzung statt) ermittelt wurden, kann auch über die Gesamteffektgröße eine Vorhersage gemacht werden. Die Effektgrößen bereits durchgeführter Untersuchungen eignen sich hierfür nur bedingt, da quantitative und qualitative Aufgaben gleichzeitig verwendet wurden; dieses Modell ist zunächst nur für quantitative Aussagen konzipiert.

Testbarkeit des Modells

Natürlich stellt sich in Anbetracht der offenen Parameter (= Häufigkeit der Prozesse, Effektgröße pro Prozeß) die Frage, ob dieses Modell überhaupt falsifizierbar ist oder ob es nicht an jede beliebige Datenlage angepaßt werden kann. Diesem Einwand kann man entgegentreten, indem man in der Modellprüfung verschiedene Wege beschreitet. Ein Weg, die Aussagekraft des Modells zu prüfen, besteht darin, die bislang nur hypothetisch angenommenen und nicht direkt beobachtbaren Prozesse in Form multinomialer Modelle (Riefer & Batchelder, 1988)¹¹ zu präzisieren und zu testen. Eine andere Vorgehensweise bestünde in der Überlegung, ob und wie die in der Literatur berichteten Befunde, insbesondere aber solche, die noch durch keine bestehende Theorie erklärt wurden („Anomalien“), in das Modell integriert

¹¹ Den Hinweis auf diese Möglichkeit, das vorliegende Modell weiterzuentwickeln und zu testen, verdanke ich Edgar Erdfelder.

werden können. Nachfolgend sind Beispiele für diese Vorgehensweise skizziert.

Bacon (1979) nimmt an, die subjektive Erinnerung, daß eine Aussage bereits früher im Experiment dargeboten worden wäre, und nicht die tatsächliche Wiederholung der Aussagen bedinge den Frequency-Validity-Effekt. Empirisch zeigte sich tatsächlich kein Unterschied in den mittleren Konfidenz-Urteilen von korrekt und fälschlich als „alt“ erinnerten Aussagen. Im Vergleich zu den als „alt“ erinnerten war das mittlere Konfidenz-Urteil von fälschlicherweise als „neu“ erinnerten Aussagen aber geringer. Eine Erklärung dieses Befundes, insbesondere aber des Urteilsverhaltens gegenüber den kritischen, d.h. den fälschlicherweise als „alt“ bzw. „neu“ klassifizierten Aussagen, sähe im Rahmen des Modells folgendermaßen aus: Obgleich im experimentellen Kontext „neu“, kann eine Aussage fälschlicherweise als „alt“ erachtet werden, wenn sie der Person aus dem präexperimentellen Kontext vertraut ist. Umgekehrt gilt, daß eine Aussage, die im Experiment bereits einmal dargeboten wurde, dann fälschlicherweise als „neu“ erachtet werden kann, wenn sie der Versuchsperson auch aus dem präexperimentellen Kontext unvertraut ist. Beurteilen die Versuchspersonen nach dem Wiedererkennungs-Urteil den Wahrheitsgehalt der Aussagen, werden bei den scheinbar „alten“ Aussagen neben anderen Cues auch der Vertrautheits-Cue angewendet, weil bei diesen Aussagen präexperimentelle Vertrautheit besteht. Bei den scheinbar „neuen“ Aussagen wird der Cue nicht aktiviert, weil diese laut obiger Annahme im präexperimentellen Kontext unvertraut sind. Die Aktivierung des Vertrautheits-Cues führt bei den fälschlicherweise als „alt“ eingeschätzten Aussagen im Mittel zu höheren Konfidenz-Urteilen. Diese Erklärung impliziert die prüfbare Hypothese, daß bei den fälschlich als „alt“ klassifizierten Aussagen präexperimentelle Vertrautheit besteht, fälschlicherweise als „neu“ klassifizierte Aussagen hingegen präexperimentell unvertraut oder unvertrauter sind.

In den Experimenten zum Hindsight-Bias, die die Tiefe der Gedächtnisspur des ursprünglichen Urteils oder der Information variierten, deutete sich folgender Zusammenhang an: Je tiefer das ursprüngliche Urteil enkodiert wurde, desto schwächer war der Hindsight-Bias ausgeprägt. Und umgekehrt, je tiefer die Information enkodiert wurde, desto größer war der Hindsight-Bias ausgeprägt. Das vorgestellte Modell würde den Zusammenhang zwischen Größe des Bias und Enkodierungstiefe durch den Prozeß „direkte Erinnerungen“ erklären. Die tiefere Enkodierung der ursprünglichen Antwort ermöglicht es, daß die ursprünglichen Urteile häufiger direkt erinnert werden können. Die tiefere Enkodierung der Information führt infolge retroaktiver Inferenz (siehe für einen Überblick Postman & Underwood, 1973) dazu, daß die ursprünglichen Urteile seltener direkt erinnert werden können. Überträgt man diese Erklärung auf die Untersuchung von Wood (1978), in der die Darbietung der Information mehrmals wiederholt wurde, dann müßte für diese Untersuchung gelten: Die Häufigkeit direkt erinnelter Urteile müßte im Vergleich zu einem

traditionellen Design, in dem die Information nur einmal dargeboten wird, geringer sein.

In der Literatur zum Hindsight-Bias wird häufig berichtet, daß der Effekt bei falschen Ereignissen in der Regel kleiner ist als bei richtigen (siehe Hawkins & Hastie, 1990; Christensen-Szalanski & Willham, 1991). Meines Wissens gibt es keine überzeugende Erklärung für diesen asymmetrischen Hindsight-Bias (siehe z.B. Fischhoff, 1977, S. 357). In dem vorliegenden Modell könnte die Asymmetrie erklärt werden. Überblickt man in Tabelle 2 die Vorhersagen über die Veränderungen des Erinnerungs-Urteils im Vergleich zum Konfidenz-Urteil, dann wird, die relative Häufigkeit der Prozesse vernachlässigt, offensichtlich, daß über alle Prozesse hinweg betrachtet, ein größerer Hindsight-Bias-Effekt für wahre als für falsche Aussagen zu erwarten ist. Dies deshalb, weil bei falschen Aussagen drei Prozesse – Generierung neuer Cues und Fehlattributionen der Informationquelle auf der einen Seite und Fehlattributionen beim Vertrauheits-Cue auf der anderen Seite – antagonistisch wirken, wohingegen bei wahren Aussagen alle Prozesse in die gleiche Richtung wirken. Die restlichen Prozesse sind hier vernachlässigbar, weil sie identische Konfidenz- und Erinnerungs-Urteile vorhersagen.

Natürlich erhebt das Modell, so wie es hier skizziert wurde, keineswegs einen Anspruch auf vollständige Beschreibung der Effekte. Im Verlauf der empirischen Überprüfung werden sicherlich manche Prozesse reformuliert, neue Prozesse aufgenommen und wieder andere völlig verworfen werden. In diesem Sinne ist das Modell ein noch vorläufiger Vorschlag, über alte Effekte neu nachzudenken.

Literatur

- ARKES, H.R., FAUST, D., GUILMETTE, T.J. & HART, K. (1988). Eliminating the hindsight bias. *Journal of Applied Psychology*, 73, 305-307.
- ARKES, H.R., HECKETT, R. & BOEHM, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2, 81-94
- BACON, F.T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 241-252.
- BEGG, I., ARMOUR, V. & KERR, T. (1985). On believing what we remember. *Canadian Journal of Behavioural Science*, 17, 199-214.
- BEGG, I., MAXWELL, D., MITTERER, J.O. & HARRIS, G. (1986). Estimates of frequency: Attribute or attribution? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 496-508.
- BIRNBAUM, M.H. (1981). Thinking and feeling: A skeptical review. *American Psychologist*, 36, 99-101.

- BIRNBAUM, M.H. & MELLERS, B.A. (1979a). One-mediator model of exposure effects is still viable. *Journal of Personality and Social Psychology*, 37, 1090-1096.
- BIRNBAUM, M.H. & MELLERS, B.A. (1979b). Stimulus recognition may mediate exposure effects. *Journal of Personality and Social Psychology*, 37, 391-394.
- BORNSTEIN, R.F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, 106, 265-89.
- BROOKS, J.O. III & WATKINS, M.J. (1989). Recognition memory and the mere exposure effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 968-976.
- CAMPBELL, J.D. & TESSER, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. *Journal of Personality*, 51, 605-620.
- CHRISTENSEN-SZALANSKI, J.J.J. & WILLHAM, C.F. (1991). The hindsight-bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 48, 147-168.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (second ed.). New York: Academic Press.
- DAVIES, M.F. (1987). Reduction of hindsight bias by restoration of foresight perspective: Effectiveness of foresight-encoding and hindsight-retrieval strategies. *Organizational Behavior and Human Decision Processes*, 43, 50-68.
- FISCHHOFF, B. (1975). Hindsight \neq Foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288-299.
- FISCHHOFF, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 349-358.
- FISCHHOFF, B. (1980). For those condemned to study the past: Reflections on historical judgment. In R. A. SHWEDER & D. W. FISKE (Eds.), *New directions for methodology of behavioral science: Fallible judgment in behavioral research*. San Francisco: Jossey-Bass.
- FISCHHOFF, B. & BEYTH, R. (1975). „I knew it would happen“. Remembered probabilities of once-future things. *Organizational Behavior and Human Performance*, 13, 1-16.
- GIGERENZER, G. (1984). External validity of laboratory experiments: The frequency-validity relationship. *American Journal of Psychology*, 97, 185-195.
- GIGERENZER, G., HOFFRAGE, U. & KLEINBÖLTING, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506-528.
- GREENE, R.L. (1984). Incidental learning of event frequency. *Memory & Cognition*, 12, 90-95.
- GREENE, R.L. (1986). Effects of intentionality and strategy on memory for frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 489-495.
- HARRIS, G., BEGG, I. & MITTERER, J. (1980). On the relation between frequency estimates and recognition memory. *Memory & Cognition*, 8, 99-104.

- HASHER, L., ATTIG, M.S. & ALBA, J.W. (1981). I knew it all along: Or, did I? *Journal of Verbal Learning and Verbal Behavior*, 20, 86-96.
- HASHER, L. & CHROMIAK, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior*, 16, 173-184.
- HASHER, L., GOLDSTEIN, D. & TOPPINO, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107-112.
- HASHER, L. & ZACKS, R.T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- HASHER, L. & ZACKS, R.T. (1984). Automatic processing of fundamental information: The case of frequency occurrence. *American Psychologist*, 39, 1372-1388.
- HAWKINS, S.A. & HASTIE, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, 107, 311-327.
- HELL, W., GIGERENZER, GAUGGEL, S., MALL, M., & MÜLLER, M. (1988). Hindsight bias: An interaction of automatic and motivational factors? *Memory & Cognition*, 16, 533-538.
- HINTZMAN, D.L. (1976). Repetition and memory. In G.H. BOWER (Hrsg.), *The psychology of learning and motivation* (Vol. 11). New York: Academic Press.
- HINTZMAN, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95, 528-551.
- HINTZMAN, D.L. & BLOCK, R.A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, 88, 297-306.
- KLAYMAN, J. & HA, Y. (1987). Confirmation, disconfirmation, and information in hypothesis-testing. *Psychological Review*, 94, 211-228.
- LINDSAY, D.S. & JOHNSON, M.K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, 17, 349-358.
- LICHTENSTEIN, S., FISCHHOFF, B., & PHILLIPS, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. KAHNEMAN, P. SLOVIC, & A. TVERSKY (Hrsg.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- MAKI, R.H. & OSTBY, R.S. (1987). Effects of level of processing and rehearsal on frequency judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 151-163.
- MORELAND, R.L. & ZAJONC, R.B. (1977). Is stimulus recognition a necessary condition for the occurrence of exposure effects. *Journal of Personality and Social Psychology*, 35, 191-199.
- MORELAND, R.L. & ZAJONC, R.B. (1979). Exposure effects may not depend on stimulus recognition. *Journal of Personality and Social Psychology*, 37, 1085-1089.
- NAVEH-BENJAMIN, M. & JONIDES, J. (1986). On the automaticity of frequency coding: Effects of competing task load, encoding strategy, and intention.

- Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 378-386.
- O'CONNOR, M. (1989). Models of human behavior and confidence in judgment: A review. *International Journal of Forecasting*, 5, 159-169.
- PENNINGTON, D.C. (1981). The British firemen's strike of 1977/78: An investigation of judgements in foresight and hindsight. *British Journal of Social Psychology*, 20, 89-96.
- POSNER, M.I. & SNYDER, C.R.R. (1975). Attention and cognitive control. In: R. L. SOLSO (Hrsg.), *Information processing and cognition: The Loyala Symposium*. Hillsdale, NJ: Erlbaum.
- POSTMAN, L. & UNDERWOOD, B.J. (1973). Critical issues in interference theory. *Memory & Cognition*, 1, 19-40.
- RIEFER, D.M. & BATCHELDER, W.H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95, 318-339.
- ROSE, R.J. & ROWE, E.J. (1976). Effects of orienting tasks and spacing of repetitions on frequency judgments. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 142-152.
- RUSSELL, J.A. & WOUZDIA, L. (1986). Affective judgments, common sense, and Zajonc's thesis of independence. *Motivation and Emotion*, 10, 169-183.
- SCHNEIDER, W. & SHIFFRIN, R.M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, 84, 1-66.
- SCHWARTZ, M. (1982). Repetition and rated truth value of statements. *American Journal of Psychology*, 95, 393-407.
- UNDERWOOD, B.J. (1969). Attributes of memory. *Psychological Review*, 76, 559-573.
- UNDERWOOD, B.J., ZIMMERMAN, J. & FREUND, J.S. (1971). Retention of frequency information with observations on recognition and recall. *Journal of Experimental Psychology*, 87, 149-162.
- WOOD, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 345-353.
- ZACKS, R.T., HASHER, L. & SANFT, H. (1982). Automatic encoding of event frequency: Further findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 106-116.
- ZAJONC, R.B. (1981) A one-factor mind about mind and emotion. *American Psychologist*, 36, 102-103.