

Eingeschränkte und ökologische Rationalität: Ein Forschungsprogramm

Ralph Hertwig und Ulrich Hoffrage

Zusammenfassung. Was ist rationales Urteilen und Entscheiden? Eine der klassischen Antworten auf diese Frage ist, dass Urteile und Entscheidungen dann rational sind, wenn sie mit den Regeln diverser normativer Systeme wie zum Beispiel der Wahrscheinlichkeitstheorie oder der „expected utility“-Theorie übereinstimmen. Mit dieser Auffassung von Rationalität geht die Fiktion einher, dass der rationale Agent über unbegrenzte Ressourcen an Zeit, Wissen und Verarbeitungskapazität verfüge. Uns Menschen stehen diese Ressourcen aber nur begrenzt zur Verfügung. Aus diesem Grund hat Herbert Simon menschliche Rationalität als eingeschränkte („bounded“) Rationalität konzipiert. Eingeschränkt rationales Urteilen ist aber keineswegs mit schlechtem Urteilen gleichzusetzen. Wir zeigen exemplarisch, dass einfache Heuristiken, die wenig Information benötigen, dennoch zu erstaunlich genauen Urteilen gelangen können. Der Schlüssel zu ihrem Erfolg liegt in ihrer ökologischen Rationalität, das heißt in ihrer Anpassung an die Struktur der Information in der Umgebung, in der sie arbeiten.

Schlüsselwörter: Eingeschränkte Rationalität, Heuristiken, Urteilen und Entscheiden, „adaptive tool box“

Bounded and ecological rationality: a research program

Abstract. What is rational judgment and decision making? One of the classical answers to this question has been that judgments and decisions are rational if they conform to the rules of various normative systems such as the probability theory or expected utility theory. Embedded within this conception of rationality is the fiction that rational agents have unlimited time, knowledge, and computational resources for making decisions. However, humans only have limited resources, which is why Herbert Simon conceptualized human rationality as „bounded.“ Boundedly rational decision making does not by any means entail poor performance. We provide examples of computationally simple heuristics that require little information, yet lead to surprisingly accurate judgments. The key to their success is ecological rationality, that is, their adaptation to the information structure of the environments in which they work.

Key words: bounded rationality, heuristics, judgment and decision making, adaptive tool box

We may look into that window [on the mind] as through a glass darkly, but what we are beginning to discern there looks very much like a reflection of the world.

Roger Shepard (1990, S. 213)

Der Herzinfarkt gehört zu den häufigsten Todesursachen in der westlichen Welt. Klinische Verfahren, die bei seiner Behandlung zum Einsatz kommen, zählen zu den Hauptursachen für die explodierenden Kosten im Gesundheitswesen der industrialisierten Länder. Auf der Webseite der *American Heart Association* findet man zum Beispiel die Schätzung, dass in den USA 1996 etwa 598.000 Bypass-Operationen durchgeführt wurden – geschätzte Kosten pro Operation \$ 44.820. Nicht jede

Person, die einen Herzinfarkt erleidet, benötigt aber den gleichen Aufwand an personeller und apparativer Versorgung. In dem Maße, in dem die Medizin effiziente, aber extrem kostenintensive Verfahren entwickelt, wird es immer wichtiger, besonders jene Personen zu identifizieren, die dieser Verfahren am dringendsten bedürfen. Wie aber soll man diese Personen identifizieren?

Breiman und Mitarbeiter (1993) haben einen Entscheidungsbaum vorgeschlagen, der zum Beispiel Herzinfarkt-Patienten auf der Grundlage von maximal drei Informationen einer „low risk“- oder „high risk“-Gruppe zuordnet: Ist der systolische Blutdruck eines Patienten unter 91, dann wird er sofort und ohne weitere Information als „high risk“ klassifiziert. Ist der Blutdruck höher und ist der Patient jünger als 62,5 Jahre, dann wird er als „low risk“ klassifiziert. Ist er allerdings älter als 62,5 Jahre, dann, und nur dann, wird eine dritte diagnostische Information erhoben: Liegt Sinus-Tachykardie vor, dann erhält der Patient den „high risk“-Status.

Wir danken Jörg Rieskamp, Hans Westmeyer, Peter Wittek und zwei anonymen Gutachtern für hilfreiche Kommentare und der DFG für finanzielle Unterstützung (Ho 1847/1-2).

Diese Entscheidungsstrategie ist in mehrfacher Hinsicht extrem einfach: Sie benötigt Informationen über maximal drei Prädiktoren – die große Mehrheit der anderen Prädiktoren bleibt unberücksichtigt (üblicherweise werden in den ersten 24 Stunden nach der Einlieferung ungefähr 100 Variablen „gescreent“; Breiman et al., 1993, S. 177). Sie stellt nur Fragen, die mit Ja oder Nein zu beantworten sind, und berücksichtigt folglich keine quantitativen Unterschiede (z. B. spielt es keine Rolle, ob jemand 23 oder 55 Jahre alt ist). Die Prädiktoren werden sequentiell abgefragt, und der Prozess kann bereits nach nur einer Frage zu Ende sein (falls der systolische Blutdruck höher als 91 ist). Sollten zwei oder drei Fragen nötig sein, dann werden die sukzessiven „Ja-Nein“-Antworten nicht integriert (z. B. in irgendeiner Weise gewichtet und addiert).

Gemessen an dem Ideal eines vollständig informierten und gründlich abgewogenen Urteils erscheint dieser Entscheidungsbaum geradezu fahrlässig. Natürlich drängt sich die Frage auf, ob diese Strategie nicht diagnostische Zielgenauigkeit für Simplizität opfert. Die erstaunliche Antwort ist: Nein. Im Gegenteil, dieser Entscheidungsbaum klassifiziert Patienten sogar besser als einige überaus komplizierte statistische Verfahren (Breiman et al., 1993). Dieses unerwartete Resultat steht für eine Beobachtung, die uns wie ein roter Faden durch diesen Artikel begleiten wird: Einfache Strategien, die lediglich ein Minimum an Information benötigen, das heißt, frugal sind, und diese Information einfach verarbeiten und dadurch Entscheidungen schnell treffen können, erlauben nichtsdestotrotz vergleichsweise akkurate Entscheidungen. Diese Beobachtung ist auch deshalb so überraschend, weil sie im Widerspruch steht zu den etablierten Grenzziehungen in den hitzigen Rationalitätsdebatten in der Psychologie, Ökonomie und Philosophie.

Eingeschränkte und ökologische Rationalität

In diesen Debatten entzündet sich stets neuer Streit an den Fragen, was eigentlich rationales Entscheiden und Urteilen unter Unsicherheit ausmacht und ob Menschen im Sinne dieser Kriterien rational sind. In den Antworten kann man – in Variationen – immer wieder zwei Auffassungen von menschlicher Entscheidungsfindung begegnen.

Die erste findet ihren vielleicht reinsten Ausdruck in Bayesianischen Entscheidungstheorien und ökonomischen „rational choice“-Theorien. Diese Theorien, denen zufolge menschliche Entscheidungen durch die Axiome normativer Theorien, insbesondere der Wahrscheinlichkeitstheorie und der „expected utility“-Theorie hinreichend gut approximiert werden können (siehe Hertwig, 1998), implizieren allerdings die Fiktion, dass Wissen, Zeit und rechnerische Kapazität unbeschränkt sind. Die zweite Auffassung nimmt die Beschränktheit menschlicher Informationsverarbeitungskapazität zum Ausgangspunkt und argumentiert, dass Menschen aufgrund

dieser Beschränkungen einfache Entscheidungsstrategien, sogenannte *Heuristiken*, verwenden. Der Preis dieser einfachen Heuristiken (z. B. Verfügbarkeitsheuristik) besteht allerdings in Urteilen und Entscheidungen, die systematisch Wahrscheinlichkeitsgesetze verletzen können (Kahneman, Slovic & Tversky, 1982). Seit Beginn der 1970er Jahre war die psychologische Urteils- und Entscheidungsforschung fast ausschließlich damit beschäftigt, diese Verletzungen zu dokumentieren.

Wie soll man es halten in dieser Debatte? Müssen wir uns zu Gunsten einer dieser beiden Positionen entscheiden? Unsere Antwort lautet „Nein“. Nein deshalb, weil es eine weitere Alternative gibt, die dort ansetzt, wo diese beiden Positionen zu kurz greifen. Sie greifen zu kurz, indem sie ignorieren, dass sich kognitive Strategien entwickelt haben können, um spezifische Ziele in Umgebungen mit spezifischen Informationsstrukturen zu erreichen.

Entscheidungsstrategien und Umweltstrukturen als ein aufeinander eingespieltes Tandem zu verstehen, ist der Kern des von Herbert Simon vorgeschlagenen Konzepts „eingeschränkte Rationalität“ (*bounded rationality*).¹ In seiner berühmten Scheren-Metapher charakterisierte Simon (1990, S. 7) rationales Verhalten von Menschen wie folgt: Es ist „shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor“. Eingeschränkte Rationalität hat also zwei verzahnte Komponenten: die Grenzen menschlicher Verarbeitungsfähigkeit und die Informationsstrukturen der Umgebungen, in denen Menschen Urteile und Entscheidungen treffen müssen. Was ist damit gemeint?

Die erste Komponente fordert, dass jegliches Modell menschlichen Denkens auf realistischen Annahmen über menschliche Verarbeitungskapazitäten basieren muss, um psychologisch plausibel zu sein. Begrenztes Gedächtnis, unvollständiges Wissen und serielle Verarbeitung sind unleugbare Konstituenten menschlicher Rationalität. Menschen sind nicht kognitiv omnipotent, und optimale Strategien sind im Regelfall ein unerreichbares Ideal. Nach Simon bleibt daher nur eins – „use approximate methods to handle most tasks“ (1990, S. 6).

Die zweite Komponente der Scheren-Metapher impliziert, dass man das Innere der „schwarzen Box“ menschlicher Kognitionen erhellen kann, indem man konsequent den Blick nach außen auf die Informationsstrukturen richtet. Worum handelt es sich bei diesen Strukturen? In „Rational choice and the structure of the environment“ verwendete Simon (1956) das Beispiel von zwei Umwelten: in einer ist die Nahrung zufällig verteilt, in der anderen existieren „Cues“, die Hinweise auf die Verteilung der Nahrung geben. Ein Organismus, der sich in der

¹ Obgleich wir damit nicht völlig glücklich sind, übersetzen wir „bounded“ als „eingeschränkte“ Rationalität; dies ist der in der Ökonomie etablierte Terminus *technicus*.

ersten Umwelt bewegt, kann mit sehr einfachen Suchstrategien überleben. In der zweiten Umwelt profitiert ein Organismus von kognitiven Fähigkeiten, die es ihm erlauben, Cue-Nahrungs-Relationen zu lernen und auf deren Basis nach Nahrung zu suchen. „Umwelt“ impliziert also keineswegs eine erschöpfende Beschreibung der physikalischen und biologischen Umgebung, sondern bezieht sich lediglich auf jene Aspekte, die für die Ziele und Bedürfnisse des Organismus relevant sind.

Trotz bedeutsamer Ausnahmen wie zum Beispiel den Arbeiten von J. R. Anderson (1990), hat Simons Betonung der ökologischen Komponente rationalen Verhaltens relativ wenig Widerhall in einer zeitgenössischen kognitiven Psychologie gefunden, die vor allem interne Prozesse thematisiert – und gelegentlich hat Simon selbst diesen vielleicht wichtigsten Schlüssel zu rationalem Verhalten vernachlässigt (z. B. Simon, 1987, S. 266). Im Folgenden werden wir drei Heuristiken vorstellen. Sie stehen stellvertretend für ein Forschungsprogramm, das ein Ensemble von eingeschränkt rationalen und ökologisch intelligenten Entscheidungsstrategien entwickelt und testet (Gigerenzer, Todd & the ABC Research Group, 1999; Chase, Hertwig & Gigerenzer, 1998). Alle diese Heuristiken vereinen zwei scheinbar widersprüchliche Eigenschaften: psychologische Plausibilität (im Sinne von Einfachheit, Schnelligkeit) und gute Performanz. Der Schlüssel zum Verständnis für diese paradox anmutende Kombination ist ihre ökologische Rationalität.

Bausteine schneller und frugaler Heuristiken

Schnelle und frugale Heuristiken verarbeiten Information in relativ einfacher Weise („fast“) und kommen mit einem Minimum an Information aus („frugal“). Jede der hier vorzustellenden Heuristiken ist in Form eines Algorithmus beschrieben, der sich aus drei Arten von Regeln zusammensetzt: Erstens, *Suchregeln*, die präzisieren, in welcher Reihenfolge Informationen eingeholt werden – zum Beispiel können die Informationen (genauer: die Prädiktoren) zufällig gesucht werden, oder die Suche kann einer Rangordnung der Prädiktoren folgen, die deren Güte für die Entscheidung widerspiegelt. Zweitens, *Abbruchregeln*, die endlose Suche vermeiden und die festlegen, wann die Suche nach Information abgebrochen wird. Zu guter Letzt verfügt jede dieser Heuristiken über eine *Entscheidungsregel*, die bestimmt, wie auf der Grundlage der gefundenen Information eine Entscheidung getroffen wird. Es sei betont, dass all diese Regeln selbst wiederum einfach sind: So wird weder die optimale Reihenfolge der Prädiktoren bestimmt, noch wird durch komplexe und zeitaufwendige Kosten-Nutzen-Berechnungen der optimale Zeitpunkt für den Abbruch der Suche nach Prädiktoren bestimmt. Darüber hinaus vermeidet jede der Entscheidungsregeln eine rechnerisch aufwendige Gewichtung und Integration der verfügbaren Information.

Die Rekognitions-Heuristik

Kann man selbst mit einem Minimum an Information korrekte Entscheidungen und Vorhersagen treffen? Ja, und die Rekognitions-Heuristik („recognition heuristic“: Goldstein & Gigerenzer, 1999) ist ein eindrucksvoller Beleg dafür. Die Fähigkeit zur Rekognition oder Wiedererkennung, auf der diese Heuristik basiert, ist eine psychologische Adaptation: zum Beispiel das Wiedererkennen von Gesichtern und Namen (ohne dass man notwendigerweise weiß, wo man das Gesicht gesehen hat oder was der Name bezeichnet). Die Heuristik kann angewandt werden, wenn von zwei Objekten dasjenige ausgewählt werden soll, das einen höheren Wert bezüglich eines bestimmten Kriteriums aufweist (das Objekt, das größer, sicherer, stärker oder dergleichen ist). Sie ist erfolgreich in jenen Situationen, in denen Rekognition mit diesem Kriterium korreliert ist. Die Rekognitions-Heuristik besagt: *Wenn von zwei Objekten eines wieder erkannt wird und das andere nicht, dann ziehe den Schluss, dass das wieder erkannte Objekt den höheren Wert hat.*² Ein Beispiel soll diese Heuristik illustrieren:

Goldstein und Gigerenzer (1999) fragten Studenten aus Deutschland und den USA: „Welche Stadt hat mehr Einwohner: San Diego oder San Antonio?“ Von den amerikanischen Studenten gaben 62 Prozent die korrekte Antwort. Obgleich die deutschen Studenten viel weniger (über amerikanische Städte) wussten als die amerikanischen, gab hier jeder die korrekte Antwort. Dies war möglich, weil alle deutschen Studenten schon von San Diego, aber viele noch nie von San Antonio gehört hatten. Sie konnten daher die Rekognitions-Heuristik anwenden, wohingegen die amerikanischen Studenten hierfür nicht „ignorant“ genug waren. Die Rekognitions-Heuristik ist also nicht nur in der Lage, bei unvollständigem Wissen zu arbeiten, sie ist sogar auf ein gesundes Ausmaß an Ignoranz – in Gestalt der Unfähigkeit zumindest einige Städte wieder zu erkennen – angewiesen.

Die Rekognitions-Heuristik und die Struktur der Umgebung

Die Rekognitions-Heuristik ist ökologisch smart: Sie profitiert von der Tatsache, dass unsere Fähigkeit, etwas wieder zu erkennen, in vielen natürlichen Umgebungen mit dem Kriterium korreliert ist, das man erschließen möchte. Die Logik der Rekognitions-Heuristik ist in Abbildung 1 veranschaulicht. Ein Kriterium, das entweder der direkten Beobachtung unzugänglich ist oder in der Zukunft liegt, soll geschätzt oder vorhergesagt werden. In vielen Fällen gibt es „Mediatoren“ in der Umgebung, die direkt zugänglich sind und das Kriterium widerspiegeln (in der Regel, ohne es offen zu legen). Ein Beispiel sind die Einwohnerzahlen amerikanischer Städte, über

² Goldstein und Gigerenzer (1999) modellieren Wiedererkennung im vorliegenden Kontext als eine binäre Variable.

die die meisten von uns kein direkt abrufbares Wissen haben. Die Größe einer Stadt spiegelt sich allerdings wider in der Häufigkeit, mit der ihr Name in Zeitungen erwähnt wird. Insofern handelt es sich bei Zeitungen um Mediatoren. Weiterhin gilt: Je häufiger der Name der Stadt im Mediator genannt wird, desto wahrscheinlicher ist es, dass eine Person ihn (wieder) erkennt. Zum Beispiel wurde in einem Zeitraum von zwei Jahren San Diego in 78 Artikeln der Wochenzeitung DIE ZEIT erwähnt, wohingegen San Antonio nur in 4 Artikeln genannt wurde (Goldstein & Gigerenzer, 1999). Dank des Mediators DIE ZEIT könnte eine Person nun erschließen, welche Stadt mehr Einwohner hat.

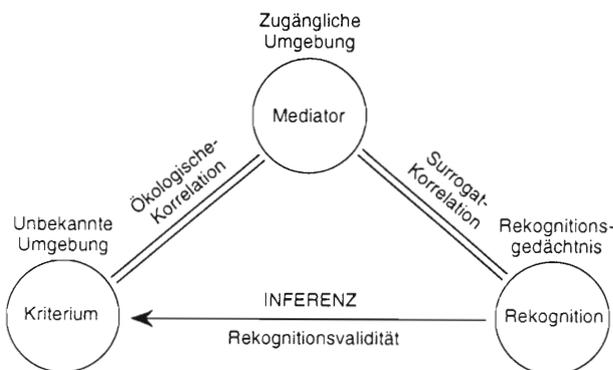


Abbildung 1. Die ökologische Rationalität der Rekognitions-Heuristik. Das nicht direkt beobachtbare Kriterium ist in der Mediator-Variablen reflektiert. Der Mediator beeinflusst die Wahrscheinlichkeit der Rekognition (Wiedererkennung). Die Schätzung des (nicht beobachtbaren) Kriteriums erfolgt auf der Grundlage der Rekognition.

Abbildung 1 zeigt die Beziehungen zwischen drei Variablen: dem Kriterium, dem Mediator und dem Rekognitionsgedächtnis (d. h. den Gedächtnisinhalten, welche der Fähigkeit, Objekte wieder zu erkennen, zu Grunde liegen). Die *ökologische Korrelation* beschreibt die Korrelation zwischen dem Kriterium (z. B. Anzahl der Einwohner) und einem Mediator (z. B. Anzahl von Artikeln in DIE ZEIT). Die *Surrogat-Korrelation* ist die Korrelation zwischen dem Mediator und dem Inhalt des Rekognitionsgedächtnisses. Die Beziehung zwischen Rekognition und dem Kriterium ist in Abbildung 1 als *Rekognitionsvalidität* (hier definiert als Korrelation) bezeichnet. Wie bereits erwähnt, haben Goldstein und Gigerenzer (1999) für einige amerikanische Städte ausgezählt, wie häufig sie in DIE ZEIT-Artikeln erwähnt worden sind. Diese Häufigkeiten haben sie dann mit der jeweiligen Anzahl der Studenten (aus einer Stichprobe der Universität Salzburg) korreliert, die den Namen dieser Städte erkannten. Diese Surrogat-Korrelation (über alle Städte) zwischen der Anzahl der Artikel und Rekognition betrug .86. Die ökologische Korrelation zwischen Anzahl der Einwohner und Artikel betrug .72, und die

Rekognitionsvalidität (d. h. die Korrelation zwischen Einwohnerzahl der Städte und der Anzahl der Studenten, die diese Städte wieder erkannt haben) betrug .66. Die Rangordnung der Koeffizienten – Surrogat-Korrelation, gefolgt von der ökologischen Korrelation und der Rekognitionsvalidität – scheint anzudeuten, dass die individuelle Rekognition stärker von der in den Medien transportierten Realität als von der physikalischen Realität geprägt ist.

Die ökologisch rationale Rekognitions-Heuristik ist eine extrem frugale Heuristik: Sie benötigt und verwendet nichts außer Rekognition. Nichtsdestotrotz funktioniert sie gut in Situationen, in denen die Rekognition mit dem Kriterium, das man erschließen möchte, korreliert ist. Auf welche Heuristiken könnte man aber zurückgreifen, wenn mehrere andere Indikatoren für das Kriterium zur Verfügung stehen und Rekognition nicht mehr anwendbar ist – wie es beispielsweise bei jenen amerikanischen Studenten der Fall war, die vor der Entscheidung „San Diego oder San Antonio“ standen?

Die Take The Best-Heuristik

Eine Heuristik, die Entscheidungen treffen kann, wenn das Rekognitionsurteil nicht diskriminiert (weil beide Städte wieder erkannt werden), ist „Take The Best“ (Gigerenzer & Goldstein, 1996). Wir erläutern die Heuristik anhand eines Problems, das für die USA und auch europäische Länder immer brennender wird: Obdachlosigkeit. In Tabelle 1 sind für vier Städte in den USA die Werte auf sechs Variablen aufgeführt, die zu den besten bekannten Prädiktoren der Obdachlosenrate in einer Stadt zählen. Die „1“ in der Spalte „Los Angeles“ und der Zeile „Mietkontrolle“ steht für die Ausübung von Mietkontrolle in dieser Stadt, und Mietkontrolle ist ein Indikator für höhere Obdachlosigkeit. Die sechs Prädiktoren sind nach ihrer Validität, das heißt nach ihrer Vorhersagegüte für Obdachlosigkeit geordnet (der Wert „1“ in der Matrix impliziert jeweils höhere Obdachlosigkeit). Wie kann man vorhersagen, welche von jeweils zwei Städten eine höhere Obdachlosenrate hat? Take The Best verwendet dazu eine schnelle, einfache Methode. Zuerst werden die Werte hinsichtlich des besten Prädiktors, Mietkontrolle, gesucht. Sind die Werte „1“ und „0“, wie beim Vergleich zwischen Los Angeles und Chicago, so wird die Suche nach weiteren Prädiktoren beendet. Auf der Grundlage des Prädiktors Mietkontrolle wird dann die Entscheidung getroffen, dass die Stadt mit der „1“, das heißt Los Angeles, die höhere Obdachlosenrate hat, was in diesem Fall korrekt ist. Sind die Werte identisch, wie beim Vergleich zwischen Chicago und New Orleans, werden die Werte der nächstbesten Prädiktoren gesucht, so lange bis ein Prädiktor gefunden ist, der zwischen den Alternativen diskriminiert.

Ebenso wie die Rekognitions-Heuristik begrenzt Take The Best die Suche nach Information. Die Abbruchregel, die Take The Best verwendet, lautet: Wenn ein Objekt (z. B. Los Angeles) einen positiven Wert („1“)

Tabelle 1. Prädiktoren für Obdachlosigkeit in US-Städten

	Los Angeles	Chicago	New York	New Orleans
Obdachlose (pro 1 Million Einwohner)	10.526	6.618	5.024	2.671
Mietkontrolle	1	0	1	0
Prozentsatz leerstehender Wohnungen ^a	1	1	1	0
Temperatur ^b	1	0	1	1
Arbeitslosigkeit ^b	1	1	1	1
Armut ^b	1	1	1	1
Sozialer Wohnungsbau ^a	1	1	0	0

Anmerkung: Die mit *a* oder *b* indizierten Prädiktoren waren ursprünglich kontinuierliche Variablen, die am Median dichotomisiert worden sind. Die Kodierung eines Prädiktors wurde so vorgenommen, dass eine „1“ immer höhere Obdachlosigkeitsraten signalisiert; *a* steht für eine positive und *b* für eine negative Beziehung des Prädiktors mit der Zielvariablen.

und das andere Objekt keinen positiven Wert hat (d. h. „0“ oder aber „unbekannt“ ist), dann wird die Suche abgebrochen. Diese einfache Abbruchregel grenzt Take The Best von der Klasse der „optimization under constraints“-Modellen ab (z. B. Anderson & Milson, 1989; Sargent, 1993; Stigler, 1961). In diesen scheinbar eingeschränkt rationalen Modellen wird eine komplexe Abbruchregel eingeschmuggelt, die die Suche nach Information *optimiert*: Die Suche nach Information wird dann abgebrochen, wenn die Kosten von weiterer Suche durch den erwarteten Nutzen zusätzlicher Information nicht mehr zu rechtfertigen sind. Aus diesem Grund sind „optimization under constraints“-Modelle schwerlich als Modelle eingeschränkter Rationalität (im Sinne von Simon) zu verstehen, obgleich sie konzipiert wurden, um der Begrenztheit menschlicher Informationsverarbeitung Rechnung zu tragen.

Ein weiteres Merkmal, das Take The Best mit der Rekognitions-Heuristik teilt, ist die einfache Entscheidungsregel: Die Entscheidung zu Gunsten eines Objekts beruht auf nur einem einzigen Prädiktor („one-reason decision making“). Man kann eine Reihe von Argumenten dafür aufführen, warum eine Person einen einzelnen guten Grund einer Kombination von Gründen vorziehen könnte. Zum Beispiel fällt es oft leichter, die wichtigste Überlegung zu identifizieren, als die Gewichte zu bestimmen, die man verschiedenen Überlegungen beimessen möchte. Die Frage, wie man verschiedene Überlegungen oder Gründe in eine Entscheidung (z. B. „heiraten“ vs. „nicht heiraten“) integriert, ist ein notorisch schwieriges Problem, das „optimal choice“-Modelle dadurch lösen, indem sie so tun, als ob es für verschiedene Überlegungen (z. B. im Alter versorgt zu sein, aber familiären Verpflichtungen nachkommen zu müssen) gemeinsame Maßeinheiten („utilities“) gäbe. Es ist aber mehr als fragwürdig, ob diese Annahme, die für eine mathematische Axiomatisierung von großer Bedeutung ist, psychologisch plausibel ist (für ein Beispiel eines Modells einge-

schränkter Rationalität, das der Inkompatibilität von Zielen Rechnung trägt, siehe die „aspiration adaptation theory“; Sauermann & Selten, 1962; Selten, 1998).

Wie gut oder schlecht ist Take The Best?

Diese Frage kann man einfach beantworten, indem man die Leistungsfähigkeit von Take The Best mit der von anderen Strategien vergleicht. Mithilfe von Computersimulationen haben Czerlinski, Gigerenzer und Goldstein (1999) Take The Best mit rechnerisch aufwendigen Standard-Algorithmen der Statistik, wie der multiplen Regressionsanalyse, verglichen. Diesen Wettstreit begann jede Strategie unter der gleichen Voraussetzung: Die spezifischen Parameter jeder Strategie wurden jeweils auf der Basis der Hälfte eines Datensatzes (Training-Set: z. B. 50% einer Menge US-amerikanischer Städte) „erlernt“. Take The Best zum Beispiel erlernte lediglich die Rangordnung der Prädiktoren (geordnet gemäß ihrer Validitäten), wohingegen multiple Regression die optimalen Gewichte (einer linearen Kombination) bestimmt. Die andere Hälfte des Datensatzes, das Test-Set, war dann die Arena, in der die Strategien gegeneinander antreten mussten.

Wie gut sind die Strategien? In Tabelle 2 findet man für zwei schnelle und frugale Strategien, Take The Best und Minimalist, sowie für multiple Regression den Prozentsatz korrekter Entscheidungen über 20 verschiedene Umgebungen. In jeder Umgebung war die Aufgabe, jeweils zwei Objekte hinsichtlich einer quantitativen Zielvariablen zu vergleichen: Zum Beispiel, welche von zwei Städten hat eine höhere Obdachlosenrate, welche von zwei Städten hat den höheren Ozongehalt, welche von zwei Personen bezieht das höhere Gehalt?

Tabelle 2 zeigt ein erstaunliches Ergebnis. Um zwischen zwei Objekten zu entscheiden, benötigte Take The

Tabelle 2. Schnelle und einfache Heuristiken (Take The Best, Minimalist) im Leistungsvergleich (über 20 Tests) mit einem rechnerisch aufwendigen statistischen Standardverfahren (multiple Regression). Bei „Vorhersage“ ist das Test-Set vom Training-Set verschieden, bei „Fitting“ sind beide Sets identisch (Czerlinski et al., 1999)

Strategie	Aufwand ^a	Leistung ^b	
		Vorhersage	Fitting
Take The Best	2,4	71	75
Minimalist	2,2	65	69
Multiple Regression	7,7	68	77

Anmerkungen: ^a Anzahl durchschnittlich verwendeter Prädiktoren.

^b Leistung entspricht der durchschnittlichen Prozentzahl korrekter Entscheidungen gemittelt über 20 Tests.

Best durchschnittlich 2,4 Prädiktoren, wohingegen multiple Regression jeweils alle verrechnete (durchschnittlich 7,7). Obwohl Take The Best damit weniger als ein Drittel der vorhandenen Information benutzte, erzielte diese einfache Heuristik im Durchschnitt mehr korrekte Vorhersagen als multiple Regression (71 % vs. 68 %)! Tabelle 2 zeigt noch eine weitere schnelle und frugale Heuristik, die mit noch weniger Information als Take The Best auskommt. Sie benutzt die gleiche Entscheidungsregel wie Take The Best, wählt aber Prädiktoren in zufälliger Reihenfolge und braucht daher noch nicht einmal deren Rangordnung zu schätzen (daher der Name „Minimalist“). Minimalist blieb nur drei Prozentpunkte hinter multipler Regression zurück.

Die rechte Spalte von Tabelle 2 zeigt als Kontrolle die Leistung bei reinem „Fitting“, bei dem die Parameter nicht aus einer Hälfte gelernt und an der anderen getestet werden, sondern bei dem Test-Set und Training-Set identisch sind. Die Differenz zwischen „Fitting“ und „Vorhersage“ reflektiert das Ausmaß, in dem eine Strategie „overfitted“, das heißt zu sehr auf die Strukturen eines Datensatzes adjustiert ist (Martignon & Hoffrage, 1999). Als Folge erweist sich die Strategie als wenig robust, sobald sie auf andere Datenstrukturen angewendet wird – sie generalisiert schlecht. Robustheit ist eine wichtige Dimension, auf der man Strategien evaluieren kann. Heuristiken, die einfach und sparsam sind, sind im Allgemeinen robuster, insbesondere in Situationen mit begrenztem Wissen. Und tatsächlich, vergleicht man die Werte in der „Vorhersage“- und der „Fitting“-Spalte, wird deutlich, dass Take The Best und Minimalist deutlich robuster sind als multiple Regression.³

Take The Best und die Struktur der Umgebung

Wie ist es möglich, dass eine so einfache Heuristik genauere Vorhersagen macht als klassische, lineare statistische Methoden, obgleich Letztere mehr Informatio-

nen benutzen und mehr Berechnungen anstellen? Ein Grund ist in der Robustheit von Take The Best zu sehen, die insbesondere in Situationen mit begrenztem Wissen („Vorhersage“-Szenario) eine gewichtige Rolle spielt. Ein anderer Grund liegt darin, dass reale Informationsstrukturen spezifische Eigenschaften haben, die einfache Heuristiken ausschöpfen können. So lässt sich zum Beispiel analytisch zeigen (Martignon & Hoffrage, 1999), dass Take The Best bei „Fitting“ dieselbe Leistung erreicht wie multiple Regression, wenn die Gewichte binärer Prädiktoren exponentiell fallend sind (z. B. 1, 1/2, 1/4, 1/8). Allgemeiner ausgedrückt: Informationen in realen Umwelten können Strukturen aufweisen, die mit einfachen Heuristiken fast genauso gut – oder sogar besser – erfasst werden wie mit komplexen statistischen Methoden. Darüber hinaus erweist sich Take The Best einem linearen Modell mit Einheitsgewichten in jenen Umgebungen überlegen, in denen die Anzahl der Prädiktoren im Vergleich zur Anzahl der Objekte klein ist (für die genauen Beziehungen siehe Martignon & Hoffrage, 1999).

³ Neben der Anzahl verwendeter Prädiktoren, der Genauigkeit und Robustheit der Inferenzen, sowie der Fähigkeit, menschliches Urteilen und Entscheiden erfolgreich zu modellieren (siehe z. B. Hoffrage, Hertwig & Gigerenzer, 2000; Rieskamp & Hoffrage, 1999), kann man Strategien auch im Hinblick auf den Aufwand evaluieren, der nötig ist, um ihre Parameter zu bestimmen (z. B. Beta-Gewichte oder Cue-Validitäten). Martignon und Laskey (1999) haben den komputationellen Aufwand verschiedener Heuristiken quantifiziert. Multiple Regression und andere komplexe Verfahren haben sich dabei als wesentlich „aufwendiger“ erwiesen als etwa Take The Best. Minimalist oder die noch folgende QuickEst-Heuristik. Dies überrascht kaum, da multiple Regression zum Beispiel das Invertieren von Interkorrelationsmatrizen erfordert, um die Beta-Gewichte festzulegen. Die Bestimmung der Rangordnung der Prädiktoren, die Take The Best benötigt, ist dazu vergleichsweise einfach. Mehr noch, wie das gute Abschneiden von Minimalist vermuten lässt, der die Cues zufällig anordnet, schneiden die Heuristiken auch dann noch relativ gut ab, wenn die Rangordnung der Prädiktoren nicht exakt den Cue-Validitäten folgt.

Die QuickEst-Heuristik

Alle bislang diskutierten Heuristiken – die Rekognitions-Heuristik, Take The Best und Minimalist – erschließen, welche von zwei Alternativen auf einer Kriteriumsvariablen die höhere quantitative Ausprägung hat. Ein solches Urteil kann man treffen, indem man lediglich eine ordinale Beziehung zwischen den zwei Alternativen herstellt. Könnte die Tatsache, dass die einfachen Heuristiken so gut mit statistisch anspruchsvollen Verfahren konkurrieren, damit zusammenhängen, dass diese Verfahren ihr Potential bei der Vorhersage einer ordinalen Beziehung nicht voll ausspielen können? Wurde hier ein Porsche auf einen holprigen Feldweg geschickt? Wenn dem so wäre, sollten die anspruchsvollen Verfahren den einfachen Heuristiken nicht dann überlegen sein, wenn zum Beispiel ein quantitatives Kriterium auf den Punkt genau vorhergesagt werden muss, wenn also – bildlich gesprochen – der Porsche wieder auf die Autobahn darf?

Diese Intuition lässt sich anhand der Leistungsfähigkeit der nächsten Heuristik testen. Die *QuickEst*-Heuristik (Kurzform für „Quick Estimation“; Hertwig, Hoffrage & Martignon, 1999) schätzt für ein gegebenes Objekt die quantitative Ausprägung auf einer Kriteriumsvariablen. Um diesen Kriteriumswert, beispielsweise die Obdachlosenrate in Chicago, zu schätzen, zieht die Heuristik wiederum Prädiktoren heran, zum Beispiel Mietkontrolle oder sozialer Wohnungsbau (siehe Tabelle 1). Wie auch Take The Best sucht QuickEst die Prädiktoren aber nicht zufällig, sondern gemäß einer bestimmten Rangordnung ab. Diese Rangordnung basiert auf dem

durchschnittlichen Kriteriumswert jener Objekte, die eine *negative* Ausprägung („0“) auf dem jeweiligen Prädiktor haben. In unserem Beispiel wird also die Rangordnung der Prädiktoren auf der Basis der durchschnittlichen Obdachlosenrate der Städte erstellt, die zum Beispiel keine Mietkontrolle oder keinen sozialen Wohnungsbau haben. Der Prädiktor mit der niedrigsten durchschnittlichen Obdachlosenrate („kleinste“ Prädiktor) erhält den ersten Rang und wird deshalb zuerst von QuickEst abgefragt.

Wie Take The Best hat auch QuickEst eine einfache Abbruchregel: Trifft die Heuristik auf einen Prädiktor, dessen Wert für das Objekt „0“ ist – Chicago zum Beispiel hat keine Mietkontrolle –, dann wird die Suche nach weiteren Prädiktoren sofort abgebrochen. Als Schätzung gibt QuickEst dann die „gerundete“ durchschnittliche Obdachlosenrate des Prädiktors an. Gerundet wird der Wert auf das Niveau der nächsten „Spontanzahl“ (Albers, im Druck). Das heißt, QuickEst muss keineswegs die durchschnittliche Obdachlosenrate des Prädiktors wissen, sondern lediglich die nächste Spontanzahl kennen. Spontanzahlen spiegeln die Tatsachen wider, dass Menschen bestimmte Zahlen des Dezimalsystems präferieren und relativ grobe Zahlen angeben, wenn sie etwas schätzen (beispielsweise 50.000 DM für den Preis des *New Beetle* anstelle von 43.295 DM).

Wie gut oder schlecht ist QuickEst?

Wie gut sind die Schätzungen von QuickEst im Vergleich zu denjenigen, die eine klassische, rechnerisch aufwendige Strategie wie zum Beispiel multiple Regression produziert? Um diese Frage zu beantworten, haben Hertwig et al. (1999) QuickEst, multiple Regression und einen Schätzbaum (der mithilfe eines aufwendigen Bayesianischen Suchprozesses als ein gutes Schätzmodell identifiziert wurde) in einem computersimulierten „Schätzwettbewerb“ gegeneinander antreten lassen. Aufgabe der Strategien war es, die Größe derjenigen deutschen Städte zu schätzen, die mehr als 100.000 Einwohner haben. Ähnlich wie in der Computersimulation mit „Take The Best“ wurde auch hier zwischen Training-Set und Test-Set unterschieden. Allerdings wurden diesmal insgesamt zehn verschiedene Größen für das Training-Set verwendet (jeweils 10 %, 20 %, 30 %, ..., 100 % aller Städte), wohingegen das Test-Set immer alle Städte beinhaltete.

Abbildung 2 zeigt den Schätzfehler der Strategien (in Form des Mittelwerts der absoluten Differenzen zwischen geschätzter und tatsächlicher Einwohnerzahl). Der Fehler ist abgetragen als eine Funktion der Größe des Training-Sets, gemittelt über jeweils 1.000 zufällige Zusammenstellungen des Training-Sets. Das 10 %-Training-Set zum Beispiel repräsentiert eine Situation, in der den Strategien für ihre Parameterschätzung (z. B. Beta-Gewichte) nur für sehr wenige Städte Wissen über die Ausprägung der Prädiktoren zur Verfügung stand. Die so aufgebauten Strategien wurden dann auf eine Umgebung (Test-Set) angewandt, die im Fall des 10 %-Training-Sets fast vollständig neu für sie war.

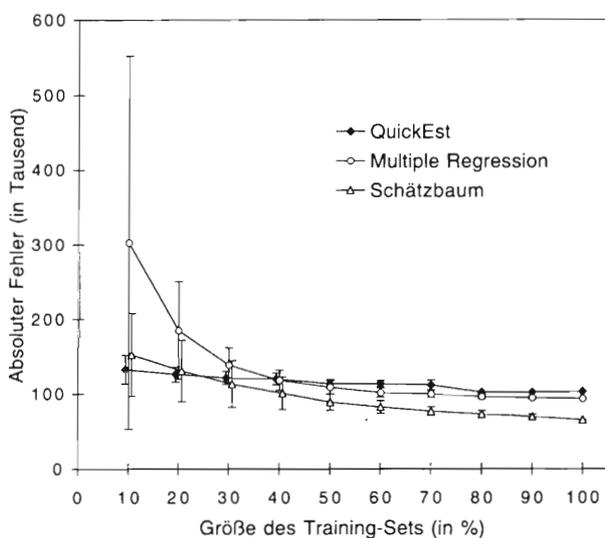


Abbildung 2. Vergleich des mittleren absoluten Fehlers (d.h. des Mittelwerts der absoluten Differenzen zwischen tatsächlichen und vorhergesagten Einwohnerzahlen), den QuickEst, multiple Regression und der Schätzbaum produzieren. Der Fehler ist als Funktion der Größe des Training-Sets abgetragen (die vertikalen Linien repräsentieren die Standardabweichung).

Wie schon Take The Best, erwies sich auch QuickEst als außerordentlich robust. QuickEst generalisierte gut auf die neue Umgebung, wohingegen zum Beispiel multiple Regression mit der neuen Umgebung relativ schlecht „zurechtkam“, das heißt bei kleinen Training-Sets relativ große Fehler machte. Wie veränderte sich die Performanz mit zunehmendem Wissen, d. h. größeren Training-Sets? Waren zum Beispiel 50 Prozent aller Städte bekannt, dann produzierten multiple Regression und QuickEst einen vergleichbar großen Schätzfehler. Aber selbst bei vollständigem Wissen (Training-Set = 100 %), ein Zustand, der psychologisch eher unplausibel ist, unterschieden sich die Schätzfehler von QuickEst (103.000) und multipler Regression (93.000) nur wenig, während der Schätzbaum hier deutlich besser abschnitt (65.000).

Die Tatsache, dass QuickEst so gut mit den komplexen Strategien konkurrieren kann und diese sogar in jenen Bereichen schlägt, in denen wenig Wissen zur Verfügung steht, ist umso bemerkenswerter, wenn man Folgendes berücksichtigt: Erstens, die Schätzungen von QuickEst können lediglich die krude Genauigkeit von Spontanzahlen erreichen, wohingegen den anderen Strategien feinabgestufte Schätzwerte zur Verfügung stehen. Zweitens, QuickEst benötigte von den insgesamt acht Prädiktoren im Schnitt nur 2,3, wohingegen der Schätzbaum 7,2 und multiple Regression alle acht Prädiktoren verwendete.⁴ Wieso zahlte QuickEst für seine psychologische Plausibilität (z. B. durch die Verwendung von Spontanzahlen) und seine Frugalität (wenig Prädiktoren) keinen höheren Preis – wieso schnitt QuickEst nicht schlechter ab?

QuickEst und die Struktur der Umgebung

Der Schlüssel zum Erfolg von QuickEst liegt wiederum in der ökologischen Intelligenz dieser Heuristik. Welche Informationsstrukturen nutzt sie aus? Trägt man die Einwohnergröße deutscher Städte (mit mehr als 100.000 Einwohnern) gegen ihre Ränge ab (wobei der Rang durch die Anzahl der Einwohner bestimmt wird, d. h., die größte Stadt erhält den ersten Rang), dann ergibt sich eine charakteristische Verteilung. In dieser Verteilung erzielen nur wenige Objekte hohe Ausprägungen, während die große Mehrzahl niedrige Ausprägungen hat (z. B. haben von allen deutschen Städten mit mehr als 100.000 Einwohnern nur 3 % über 1.000.000 Einwohner, wohingegen etwa 54 % zwischen 100.000 und 200.000 Einwohner haben). Durch diesen Typus von Verteilung, die

⁴Verwendete multiple Regression nur jene Regressionskoeffizienten, die signifikant von Null verschieden waren, dann waren es im Schnitt 7,3 Prädiktoren. Auf der Basis dieser Prädiktoren verringerte sich der Schätzfehler im 10 %-Training-Set auf 225.000 – war also immer noch deutlich größer als für QuickEst; für alle anderen Training-Sets (20 %–100 %) fiel der Schätzfehler geringfügig (im Schnitt 2.500) kleiner aus.

häufig als J-Verteilung bezeichnet wird, lassen sich eine Vielzahl von Phänomenen beschreiben: zum Beispiel die Verteilung von Erdbebenstärken, Abrufwahrscheinlichkeiten von Gedächtnisspuren und Vermögen von Haushalten (für genauere Charakterisierung von J-Verteilungen siehe Grüneis et al., 1989).

QuickEst nutzt die J-Verteilung auf zwei Arten aus: In einer Verteilung, in der die meisten Objekte geringe Ausprägungen haben, sind auch die meisten Prädiktorwerte negativ (z. B. haben von den Städten mit mehr als 100.000 Einwohnern die meisten *keinen* Flughafen, *kein* Team in der Fußball-Bundesliga usw.). Die Abbruchregel von QuickEst – „breche Suche ab, sobald ein Prädiktor mit einer negativen Ausprägung gefunden wurde“ – erlaubt es, die Suche schnell zu beenden. Da QuickEst außerdem die Prädiktoren vom „kleinsten“ zum „größten“ Prädiktor absucht, wird jedes Objekt, für das die Suche bereits nach ein oder zwei Prädiktoren abgebrochen wird, als relativ klein eingeschätzt. In J-verteilten Welten, in denen die meisten Objekte in der Tat klein sind, führt diese ökologisch rationale Heuristik zu frugalen und gleichzeitig erstaunlich akkuraten Urteilen.

Schlussbemerkungen

Intelligenz im Sinne von eingeschränkter und ökologischer Rationalität besteht nicht aus einem universellen Kalkulus, auf den sich alles Denken reduzieren ließe, sondern aus einer adaptiven Werkzeugkiste mit einer Vielzahl von Heuristiken (Gigerenzer & Selten, in press). Die hier skizzierten Heuristiken sind an spezifische Situationen gebunden; die Rekognitions-Heuristik etwa benötigt ein bestimmtes Ausmaß an Ignoranz, die QuickEst-Heuristik ist an die J-Verteilung adaptiert. Keine der Heuristiken ist universell. Man kann aber spekulieren, dass andere Heuristiken sich aus ähnlichen Bausteinen, das heißt aus einfachen Suchregeln, Suchabbruchregeln und Entscheidungsregeln zusammensetzen. Die Güte dieser Heuristiken bemisst sich nicht nach ihrer Übereinstimmung mit dem normativen Regelwerk klassischer Kanons menschlicher Rationalität, sondern nach ihrer Fähigkeit, mit begrenztem Wissen, in begrenzter Zeit und mit wenig Aufwand schnelle und zutreffende Urteile über die – physikalische und soziale – Umwelt zu treffen.

Literatur

- Albers, W. (in press). Prominence theory as a tool to model boundedly rational decisions. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 297–317). Cambridge, MA: MIT Press.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703–719.

- Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1993). *Classification and regression trees*. New York: Chapman & Hall.
- Chase, V. M., Hertwig, R. & Gigerenzer, G. (1998). Visions of rationality. *Trends in Cognitive Sciences*, 2, 206–214.
- Czerlinski, J., Gigerenzer, G. & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd & the ABC Research Group, *Simple heuristics that make us smart* (pp. 97–118). New York: Oxford University Press.
- Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–669.
- Gigerenzer, G. & Selten, R. (Eds.). (in press). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.
- Gigerenzer, G., Todd, P. M. & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Goldstein, D. G. & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. M. Todd & the ABC Research Group, *Simple heuristics that make us smart* (pp. 37–58). New York: Oxford University Press.
- Grüneis, F., Nakao, M., Yamamoto, M., Musha, T. & Nakahama, H. (1989). An interpretation of 1/f fluctuations in neural spike trains during dream sleep. *Biological Cybernetics*, 60, 161–169.
- Hertwig, R. (1998). Psychologie, experimentelle Ökonomie und die Frage, was gutes Experimentieren ist. *Zeitschrift für Experimentelle Psychologie*, 45, 2–19.
- Hertwig, R., Hoffrage, U. & Martignon, L. (1999). Quick estimation: Letting the environment do the work. In G. Gigerenzer, P. M. Todd & the ABC Research Group, *Simple heuristics that make us smart* (pp. 209–234). New York: Oxford University Press.
- Hoffrage, U., Hertwig, R. & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 566–581.
- Kahneman, D., Slovic, P. & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Martignon, L. & Hoffrage, U. (1999). Why does one reason-decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd & the ABC Research Group, *Simple heuristics that make us smart* (pp. 119–140). New York: Oxford University Press.
- Martignon, L. & Laskey, K. B. (1999). Bayesian benchmarks for fast and frugal heuristics. In G. Gigerenzer, P. M. Todd & the ABC Research Group, *Simple heuristics that make us smart* (pp. 169–188). New York: Oxford University Press.
- Rieskamp, J. & Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell? In G. Gigerenzer, P. M. Todd & the ABC Research Group, *Simple heuristics that make us smart* (pp. 141–167). New York: Oxford University Press.
- Sargent, T. J. (1993). *Bounded rationality in macroeconomics*. Oxford: Oxford University Press.
- Sauerermann, H. & Selten, R. (1962). Anspruchsanpassungstheorie der Unternehmung. *Zeitschrift für die gesamte Staatswissenschaft*, 118, 577–597.
- Selten, R. (1998). Aspiration adaptation theory. *Journal of Mathematical Psychology*, 42, 191–214.
- Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69, 213–225.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.
- Simon, H. A. (1987). Bounded rationality. In J. Eatwell et al. (Eds.), *The New Palgrave: A dictionary of economics* (pp. 266–268). London: Macmillan.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41, 1–19.
- Shepard, R. N. (1990). *Mind sights*. New York: Freeman.

Ralph Hertwig

Max-Planck-Institut für Bildungsforschung
Lentzeallee 94
D-14195 Berlin
E-Mail: hertwig@mpib-berlin.mpg.de

