

# Scientia Nova

Herausgegeben von

Rainer Hegselmann, Gebhard Kirchgässner,

Hans Lenk, Siegwart Lindenber,

Werner Raub, Thomas Voss

Bisher erschienen u. a.:

*Robert Axelrod*, Die Evolution der Kooperation

*Karl H. Borch*, Wirtschaftliches Verhalten bei Unsicherheit

*Churchman / Ackoff / Arnoff*, Operations Research

*James S. Coleman*, Grundlagen der Sozialtheorie

Erklären und Verstehen in der Wissenschaft

Evolution und Spieltheorie

*Bruno de Finetti*, Wahrscheinlichkeitstheorie

*Richard C. Jeffrey*, Logik der Entscheidungen

*Nagel / Newmann*, Der Gödelsche Beweis

*John von Neumann*, Die Rechenmaschine und das Gehirn

*Erhard Oeser*, Wissenschaft und Information

*Howard Raiffa*, Einführung in die Entscheidungstheorie

*Erwin Schrödinger*, Was ist ein Naturgesetz?

*Rudolf Schüßler*, Kooperation unter Egoisten: vier Dilemmata

*Thomas Voss*, Rationale Akteure und soziale Institutionen

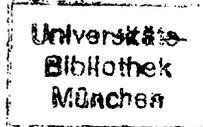
*Hermann Weyl*, Philosophie der Mathematik und Naturwissenschaft

# Moralische Entscheidung und rationale Wahl

Herausgegeben von  
Martin Hollis und Wilhelm Vossenkuhl

R. Oldenbourg Verlag München 1992

053 468 279



Die Deutsche Bibliothek - CIP-Einheitsaufnahme

**Moralische Entscheidung und rationale Wahl** / hrsg. von  
Martin Hollis und Wilhelm Vossenkuhl. - München :  
Oldenbourg, 1992  
(Scientia nova)  
ISBN 3-486-55949-4  
NE: Hollis, Martin [Hrsg.]

© 1992 R. Oldenbourg Verlag GmbH, München

Das Werk einschließlich aller Abbildungen ist urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Bearbeitung in elektronischen Systemen.

Umschlaggestaltung: Dieter Vollendorf  
Gesamtherstellung: WB-Druck, Rieden a.F.

ISBN 3-486-55949-4

92 P 486

# Inhalt

Einleitung	1
<i>Karl Homann</i> Die ökonomische Dimension von Rationalität	11
<i>Edward McClennen</i> Morality as a Public Good	25
<i>Robert Sugden</i> Contractarianism and Norms	43
<i>Wulf Gaertner</i> Einige Überlegungen zu A. Sens Formulierung individueller Rechte	63
<i>Nick Baigent</i> Deliberation and Rational Choice	75
<i>Isaac Levi</i> Consensus, Rationality and Pareto Unanimity	85
<i>Ulrich Krause</i> Rationalität angesichts von Konflikten	101
<i>Martin Hollis</i> Honour Among Thieves	115
<i>Julian Nida-Rümelin</i> Ökonomische Rationalität und praktische Vernunft	131
<i>Wilhelm Vossenkuhl</i> Vernünftige Wahl, rationale Dilemmas und moralische Konflikte	153
<i>Albert Weale</i> Political Accountability and Normative Rationality	175
<i>Rainer Hegselmann</i> Moralität im iterierten Gefangenen-Dilemma	183
<i>John Skorupski</i> Value and Distribution	191

Namensregister	209
Sachregister	215
Autorenverzeichnis	221

## Ökonomische Rationalität und praktische Vernunft

Die praktische Philosophie im weiten Sinne des Nachdenkens über menschliche Praxis war seit ihren Anfängen von zwei konkurrierenden Intuitionen geprägt. Die Intuition der Folgenoptimierung besagt, daß ein Mensch vernünftig oder rational handelt, wenn er das, was er für erstrebenswert hält, durch seine Handlungen optimiert. Vielleicht erweckt diese Formulierung den Eindruck, daß die hier gemeinte Intuition sehr präzise sei. Dies ist natürlich nicht der Fall, daher ist folgende Formulierung vorzuziehen: Eine Person handelt rational (oder vernünftig), wenn ihre Handlungen im Hinblick auf die Ziele dieser Person sinnvoll erscheinen (als ein gutes Mittel gelten können, um diese Ziele zu erreichen). Die andere Intuition (die Intuition der Prinzipienorientierung) besagt, daß vernünftiges Handeln darin besteht, bestimmten Prinzipien oder Kriterien zu gehorchen.

Erst in jüngerer Zeit werden diese beiden Intuitionen praktischer Rationalität als gegensätzlich oder sogar unvereinbar empfunden. Bei Aristoteles gehen beide Intuitionen eine komplizierte Verbindung ein, was etwa bei der Einführung des sogenannten praktischen Syllogismus deutlich wird.<sup>1</sup> Aber auch noch bei den Klassikern des Utilitarismus im Umfeld der englischen Nationalökonomie<sup>2</sup> bleiben beide Intuitionen miteinander verwoben, was heute als eine unzulässige Vermengung zweier ganz unterschiedlicher Utilitarismus-Varianten, der des Regel- und der des Handlungsutilitarismus erscheint.<sup>3</sup> Erst Kant setzt diese beiden Intuitionen in einen scharfen Gegensatz: Ethisches Handeln ist mit Folgenoptimierung unvereinbar, nur prinzipienorientiertes, dem moralischen Gesetz oder dem kategorischen Imperativ verpflichtetes, Handeln ist moralisch. Wenn seither von "praktischer Vernunft" die Rede ist, so assoziiert man ein mehr oder weniger stark kantisch geprägtes Verständnis praktischer Rationalität. Die ökonomische Theorie hat sich dagegen von Anbeginn auf eine Konzeption praktischer Rationalität gestützt, die ganz auf die Intuition der Folgenoptimierung<sup>4</sup> vertraute. Die im Umfeld der Ökonomie entstandenen Modelle rationalen Entscheidens - die Entscheidungstheorie im engeren Sinne, die Spieltheorie und die Theorie kollektiver Entscheidungen (collective choice) - haben unterdessen auch außerhalb des ökonomischen Sektors vielfältige Anwendungsmöglichkeiten gefunden und damit zu einer Dominanz des ökonomischen Paradigmas praktischer Rationalität zumindest im angelsächsischen Sprachraum beigetragen. Paradoxerweise hat jedoch gerade die Verfeinerung der entscheidungstheoretischen Analyse und ihre Übertragung auf

komplexe Interaktionssituationen und institutionelle Strukturen dazu beigetragen, das ökonomische Paradigma der Folgenoptimierung<sup>5</sup> in Frage zu stellen. Die zahlreichen Dilemmata und Paradoxa der Spieltheorie und der Theorie kollektiver Entscheidungen unterminieren das Ausgangsparadigma<sup>6</sup> selbst. Die Konsequentialismus-Kontroverse der vergangenen Jahre hat deutlich werden lassen, welche Probleme das ökonomische Paradigma in der Ethik aufwirft.<sup>7</sup> Gegenüber der Konkurrenz der beiden genannten Grundintuitionen und der auf diesen beruhenden paradigmatischen Konzeptionen praktischer Rationalität kann man vier unterschiedliche Standpunkte einnehmen:

(1) Man kann der Auffassung sein, daß diese beiden Intuitionen bzw. die entsprechenden Paradigmata nur scheinbar in einem Gegensatz zueinander stehen. Dieser Auffassung sind die meisten Neo-Aristoteliker, aber auch etwa Richard Hare.<sup>8</sup>

(2) Eine andere Möglichkeit des Ausgleichs besteht darin, bereichsspezifische Rationalitäten anzunehmen. Etwa, indem man zwischen Handlungen unterscheidet, die einer moralischen Beurteilung unterliegen und solchen, die außerhalb moralischer Beurteilungen stehen. Genuin moralisches Verhalten wäre dann prinzipienorientiert und (rationales) außermoralisches Verhalten folgenoptimierend. Es liegt auf der Hand, daß das Verhältnis dieser beiden Bereiche zueinander diffizile Probleme aufwirft, aber wenn man dieses Abgrenzungsproblem für lösbar erachtet, wäre der Konflikt dieser beiden Intuitionen nach dieser Auffassung nicht nur erklärt, sondern auch behoben.

Der dritte und vierte Standpunkt sind im Gegensatz zu den genannten 'kolonialistisch': während aber der Kolonialismus des ökonomischen Paradigmas, der Prinzipien den Status von bloßen Faustregeln zuweist oder andere Methoden wählt, um Prinzipienorientierung im folgenoptimierenden Paradigma aufgehen zu lassen (3), beschränkt sich der Kolonialismus des prinzipienorientierten Paradigmas praktischer Rationalität (4) bislang auf die Ethik und einzelne Diskussionsbeiträge zur Spieltheorie.

Meine persönliche Auffassung ist, daß - abgesehen von einem Grenzfall - praktische Rationalität prinzipienorientiert ist, was allerdings keineswegs heißt, daß die Folgen einer Handlung für ihre Beurteilung als rational keine Rolle spielen. Folgenoptimierung, allerdings nicht in einem punktuellen, sondern in einem strukturellen Sinne, scheint mir sogar im Zentrum einer adäquaten Theorie praktischer Rationalität zu stehen. Ich werde jetzt nicht versuchen, diesen Standpunkt darzustellen und zu rechtfertigen - das Ziel dieses Beitrages ist bescheidener: Ich möchte diese beiden genannten Grundintuitionen praktischer Rationalität soweit präzisieren, wie es erforderlich ist, um die auf diesen Intuitionen beruhenden paradigmatischen Modelle praktischer Rationalität einander gegenüber zu stellen. Die Argumentation ist dabei so "sparsam" wie möglich, um vom Zentrum der Kontroverse ablenkende Diskussionen zu vermeiden.<sup>9</sup> Trotz der zahlreichen Verfeinerungen, die unterdessen vorgeschlagen wurden, genügt es, als Axiomatisierung der ökonomischen Theorie

praktischer Rationalität im Sinne des revealed-preference-Konzeptes die v. Neumann-Morgensternsche Darstellung heranzuziehen und die Defizite des ökonomischen Rationalitätsmodells anhand des viel diskutierten Gefangenendilemmas aufzuzeigen. Dabei wird sich allerdings herausstellen, daß das prinzipienorientierte Modell praktischer Rationalität mit der v. Neumann-Morgensternschen Axiomatik vereinbar ist.<sup>10</sup>

### *1. Ökonomische Rationalität*

Die ökonomische Konzeption praktischer Rationalität präzisiert und spezifiziert die Intuition, eine rationale Entscheidung zeichne sich dadurch aus, geeignetes Mittel zu sein, um die angestrebten Folgen zu realisieren. Es handelt sich dabei insofern um eine spezifische Interpretation als die Theoretiker der Ökonomie eine subjektivistische gegenüber der eher objektivistischen Sicht der klassischen teleologischen Handlungskonzeption bevorzugen. Der rationale Akteur optimiert die von ihm (subjektiv) erstrebten Folgen seines Handelns. Die hierarchische Struktur des Guten, wie es etwa im ersten Buch der Nikomachischen Ethik geschildert ist, wird in dieser Sicht vom Kopf auf die Füße gestellt: Das Gute für die Polis bestimmt nicht mehr das Gute für den Einzelnen, sondern das Gute für die Polis ist die Aggregation subjektiver Einzelinteressen. Insofern der Adressat der ökonomischen Theorie die sich gerade herausbildenden Nationalstaaten waren, wurde das Kriterium rationalen staatlichen Handelns diese individualistisch bestimmte kollektive Wohlfahrtsfunktion.

Die Stringenz dieser Sichtweise, die jede Person mit gleichem Gewicht ausstattete, zu einer Zeit, als die politischen Institutionen noch weit entfernt von dem Prinzip "one man one vote" waren, machte zugleich ihren reformerischen Impetus aus, wie er sich in den politischen Aktivitäten der utilitaristischen "Gründerväter" niedergeschlagen hat. Auffallend spät - erst gegen Mitte dieses Jahrhunderts wird deutlich, daß diese Stringenz nur unter Ausblendung eines wichtigen Tatbestandes aufrechterhalten werden konnte, nämlich dem, daß auch politisches Handeln sich durch das Handeln einzelner Akteure konstituiert. Das zentrale Dogma des Wirtschaftsliberalismus, die individuelle Optimierung subjektiver Interessen optimiere auch die Aggregation dieser subjektiven Interessen (die kollektive Wohlfahrtsfunktion), läßt sich spätestens seit Mancur Olsons "Logic of Collective Action"<sup>11</sup> nicht länger aufrechterhalten, was die theoretische Harmonie der Verbindung von Liberalismus, Utilitarismus und Demokratie empfindlich stört. Die ökonomische Theorie der Politik von Buchanan bis Axelrod kann man als eine Folge unterschiedlicher Ansätze zur Wiederherstellung dieser Harmonie betrachten.

Bezeichnenderweise sind in diesen neuen Anwendungsbereichen der ökonomischen Theorie die verfeinerten Konzeptionen des revealed-preference-

Paradigmas nirgendwo zur Geltung gekommen. Die Attraktivität des ökonomischen Analyseinstrumentariums beruht nach wie vor auf der für die ökonomische Theorie konstitutiven Vorstellung, daß rationale Akteure ihren Eigennutzen optimieren. Nun gibt es gegen diesen normativen Egoismus der ökonomischen Theorie Einwände unterschiedlichster Provenienz: empirisch psychologische, die diese Annahme aufgrund von Motivationsforschungen bestreiten; anthropologische wie die der schottischen Moralphilosophie und besonders Joseph Butlers; methodologische, die abstrakte Idealtypen dieser Art für unzulässig halten; epistemologische, die die mangelnde Operationalisierbarkeit des Begriff Eigennutzen kritisieren etc. Letztgenannte haben eine wesentliche Rolle für die Entwicklung des revealed-preference-Konzeptes gespielt, das den problematischen Begriff des Eigennutzens auf den der offenbaren Präferenzen, soweit sie konsistent sind, zurückführt.

Keine dieser Kritiken wird im folgenden eine Rolle spielen. Der Grund dafür ist der folgende: alle diese Kritiken lassen sich jeweils durch eine Liberalisierung des ökonomischen Paradigmas praktischer Rationalität relativ problemlos auffangen. Diese Kritiken gefährden daher das ökonomische Paradigma nur insofern, als sie dessen weitgehende Trivialisierung zur Folge haben können, während ich plausibel machen möchte, daß das ökonomische Paradigma als Kern einer generellen Theorie praktischer Rationalität inadäquat ist.

Ich versuche dabei, die gedanklichen Eindrücke - wenn auch sehr kondensiert - nachzuzeichnen, die mich selbst zu einem Paradigmenwechsel veranlaßt haben. Vor diesem Paradigmenwechsel schienen mir alle gängigen Einwände gegen das ökonomische Modell der Folgenoptimierung durch entsprechende Modifikationen der Wertfunktion, der Rolle epistemischer Bedingungen, der Anreicherung des Entscheidungsmodells durch mehrstufige Verfahren, durch die Einführung von parametrisierter Rationalität beantwortbar zu sein. Auch das Problem der Verteilungsgerechtigkeit kann noch nicht als Argument gegen das ökonomische Modell, sondern nur gegen bestimmte radikale, d.h. ordinalistische Varianten, vorgebracht werden. Erst das Problem der Koordination individueller Entscheidungen schien mir im Rahmen des ökonomischen Paradigmas nicht mehr adäquat lösbar zu sein.

Andererseits erscheinen die Axiome des revealed-preference-Konzeptes der modernen Nutzentheorie als Merkmale der Präferenzen einer idealen rationalen Person durchaus plausibel zu sein, auch wenn die Prüfung ihrer empirischen Realisierung methodische Probleme aufwirft und die konstitutive Verknüpfung von Wünschbarkeits- und Wahrscheinlichkeitsfunktion unbefriedigend bleibt. Wenn aber die Axiome der modernen ökonomischen Nutzentheorie plausibel sind, dann scheint man sich (mit logischer Notwendigkeit) auf ein folgenoptimierendes Modell praktischer Rationalität festgelegt zu haben.

Betrachten wir die Kohärenzaxiome der ökonomischen Nutzentheorie: Sei  $X$  eine Menge von Alternativen, die von den Präferenzen einer idealen rationalen Person in eine Rangfolge gebracht wird, sei  $R$  die Relation der schwachen

Präferenz, d.h.  $xRy$  heißt nichts anderes, als "die betreffende Person zieht  $x$  gegenüber  $y$  vor, oder ist indifferent zwischen  $x$  und  $y$ " bzw. "die betreffende Person findet  $x$  mindestens ebenso gut wie  $y$ ".

*1. Reflexivität:*

Die (ideale) rationale Person hat eine schwache Präferenz von  $x$  gegenüber  $y$ , das gilt für alle Alternativen aus  $X$ :

*2. Vollständigkeit:*

Die (ideale) rationale Person hat eine schwache Präferenz für  $x$  gegenüber  $y$  oder eine schwache Präferenz von  $y$  gegenüber  $x$ , das gilt für alle Elemente  $x$ ,  $y$  aus  $X$ .

*3. Transitivität:*

Die (ideale) rationale Person hat eine schwache Präferenz für  $x$  gegenüber  $z$ , wenn sie eine schwache Präferenz von  $x$  gegenüber  $y$  hat und zugleich eine schwache Präferenz für  $y$  gegenüber  $z$ .

Man kann diese Axiome als Bedeutungspostulate des Begriffs Präferenz interpretieren, aber es ist naheliegender, zumindest die beiden letzten Axiome als Rationalitätspostulate, d.h. als notwendige Charakteristika der Präferenzen einer rationalen Person zu verstehen.

Diese drei Axiome reichen jedoch noch nicht hin, um vom qualitativen Begriff der Präferenz zum quantitativen Begriff einer subjektiven Wertfunktion überzugehen, die der rationale Akteur maximiert. Eine quantitative Funktion wäre allein aufgrund dieser drei Axiome noch unterbestimmt, d.h. jede monotone Transformation einer Funktion, die mit einer diese drei Axiome erfüllenden Präferenzrelation kompatibel ist, wäre ebenfalls kompatibel. Um von einer Ordinal- zu einer Kardinalskala überzugehen, ist es erforderlich, die Alternativenmenge probabilistisch zu verfeinern. Die Präferenzen der idealen rationalen Person sollen nicht nur die (festen) Alternativen aus  $X$ , sondern auch die Wahrscheinlichkeitsverteilungen (Lotterien) über  $X$  in eine konsistente Rangordnung bringen. Auch bezüglich der probabilistisch erweiterten Alternativenmenge  $X^*$  sind die Präferenzen einer rationalen Person reflexiv, vollständig und transitiv, zusätzlich werden folgende vier Axiome postuliert:

*Reduktion;*

Die (ideale) rationale Person ist indifferent zwischen zwei Wahrscheinlichkeitsverteilungen über  $X$ , wenn eine in die andere durch Umformungen gemäß dem wahrscheinlichkeitstheoretischen Kalkül überführbar ist.

*Substitution:*

Wenn eine (ideale) rationale Person zwischen einer Wahrscheinlichkeitsverteilung  $x^*$  und einer bestimmten Alternative  $x$  aus  $X$  indifferent ist, dann

kann  $x^*$  und  $x$  in beliebigen Kontexten ausgetauscht werden, ohne daß sich die Präferenzen ändern.

*Kontinuität:*

Wenn  $x_a$  die beste und  $x_z$  die schlechteste Alternative in  $X$  ist, dann gibt es für jede Alternative  $x$  aus  $X$  eine Wahrscheinlichkeit  $p$  für  $x_a$ , so daß die (ideale) rationale Person zwischen der Verteilung  $[px_a \text{ \& } (1-p)x_z]$  und  $x$  indifferent ist.

*Monotonie:*

Von zwei Wahrscheinlichkeitsverteilungen zwischen  $x$  und  $y$  zieht die (ideale) rationale Person diejenige vor, die eine höhere Wahrscheinlichkeit für die bessere Alternative vorsieht.

Die bisher genannten Axiome sind allesamt - zumindest prima facie - intuitiv plausibel. Keines dieser Axiome nimmt inhaltlich auf bestimmte Motivationen der handelnden Person Bezug oder macht Voraussetzungen darüber, was für die betreffende Person subjektiv wertvoll sei. Die Plausibilität dieser Axiome hängt also nicht von einer bestimmten inhaltlichen Festlegung etwa auf ökonomische Ziele ab. Die Plausibilität dieser Alternativen scheint auch unabhängig davon zu sein, wie man die Alternativenmenge  $X$  interpretiert. Mögen nun die Elemente von  $X$  monetäre Auszahlungen an den Akteur oder mögliche Welten sein, es ist prima facie kein Grund erkennbar, der zur Aufgabe eines dieser Axiome veranlassen könnte. Was immer der geeignete Gegenstand der wertenden Beurteilung oder der im Verhalten sich manifestierenden Präferenzen einer Person sein mag, die Axiome scheinen nicht mehr zu sein als Minimalbedingungen, die an die praktische Rationalität dieser Person gestellt werden müssen.

Wenn diese obengenannten Axiome simultan von den Präferenzen einer Person erfüllt sind, dann läßt sich eine reellwertige Funktion über  $X$  angeben, die von den Präferenzen dieser Person maximiert wird, d.h. wann immer die Person  $x$  gegenüber  $y$  präferiert, dann ordnet diese Funktion  $x$  einen höheren Zahlenwert zu als  $y$ . Eine Funktion dieser Art wird in der ökonomischen Theorie etwas voreilig "Nutzenfunktion" genannt, sie ist linear und bis auf lineare Transformation aufgrund einer im obigen Sinne kohärenten Präferenzrelation eindeutig festgelegt.

Da dieses Theorem aus den oben normalsprachlich wiedergegebenen Axiomen logisch-deduktiv ableitbar ist, ist jeder, der diese Axiome als adäquate Merkmale einer Präferenzrelation einer rationalen Person ansieht (logisch) gezwungen, auch anzuerkennen, daß das Entscheidungsverhalten der betreffenden Person diese reell-wertige Funktion über  $X$  maximiert, sofern er bereit ist, zwischen Entscheidungsverhalten und Präferenzen jene direkte Korrelation herzustellen, wie sie dem revealed-preference-Konzept entspricht. Da also gezeigt ist, daß eine quantitative Funktion optimiert wird, liegt es nahe,

diese Funktion als die quantitative Repräsentation subjektiver Wünsche und Motive, bzw. als jene Nutzenfunktion zu interpretieren, die die klassische Ökonomie vergeblich zu operationalisieren versucht hat.

Eine Entscheidung wäre damit rational, wenn sie die motivierenden Intentionen der handelnden Person (repräsentiert durch die quantitative Funktion) optimiert. Die einzelnen Zustände (die Elemente aus  $X$ ) realisieren die motivierenden Intentionen der handelnden Person in unterschiedlichem Maße und die Rangordnung dieser Zustände bemißt sich nach dem Maß dieser Erfüllung. Das ökonomische Rationalitätsmodell im engeren Sinne würde dieses Maß mit dem persönlichen Wohlergehen des Akteurs identifizieren, während die ökonomische Rationalitätskonzeption im weiteren Sinne nur auf dem instrumentellen Charakter der Entscheidung für die Präferenzen-Erfüllung beharrt, aber zuläßt, daß die betreffenden motivierenden Intentionen nicht-egoistischer Natur sind.

Weder die enge noch die weite Fassung werden jedoch durch die oben angeführten Kohärenzpostulate impliziert. Bei der Formulierung der Kohärenzpostulate wurde an keiner Stelle auf motivierende Intentionen Bezug genommen. Diese Postulate schränken die Gestalt der Präferenzrelation ein, aber sie beinhalten keinerlei Annahmen über die Art und Weise des Zustandekommens dieser Präferenzen.

Es ist - zumindest prima facie - nicht ausgeschlossen, daß der ideale kantische Akteur, dessen Handlungsmaximen dem moralischen Gesetz gehorchen, eine Präferenzrelation bildet, die alle Kohärenzpostulate erfüllt. Wenn dem so ist, dann gibt es - wie das Nutzentheorem zeigt - eine reellwertige Funktion, die dieser kantische Akteur maximiert. Aber zugleich wäre es ausgeschlossen, daß dieser Akteur durch die Maximierung dieser Wertfunktion motiviert ist. Sollte diese Vereinbarkeit bestehen, dann wäre die dem kantischen Akteur zugeordnete Nutzenfunktion nur formal, d.h. sie würde keine motivierenden Intentionen repräsentieren und damit wäre in diesem Fall die zentrale Idee der ökonomischen Rationalitätskonzeption, eine rationale Handlung optimiere bestimmte subjektive Ziele, und diese Ziele realisierten sich durch die Konsequenzen der Handlung, nicht erfüllt. Ein Vertreter einer kantischen Rationalitätskonzeption scheint in keiner Weise gezwungen zu sein, die Kohärenzpostulate zurückzuweisen, auch wenn er die ökonomische Vorstellung von Rationalität auch in ihren liberaleren Varianten für falsch hält. Insofern konstituiert sich das, was man unter ökonomischer Rationalität versteht, nicht in den Kohärenzpostulaten, sondern in einer ganz spezifischen Interpretation der durch sie konstituierten quantitativen Funktionen. Wie wir in den folgenden Abschnitten sehen werden, bestätigt sich diese Vermutung bei einer etwas sorgfältigeren Betrachtung des Verhältnisses von praktischer Vernunft und ökonomischer Rationalität.

## 2. Praktische Vernunft

Von alters her stehen zwei Fragen im Zentrum des ethischen Diskurses:

1. Was ist wertvoll, was ist erstrebenswert, welche Lebensform ist gut?
2. Angenommen, man weiß, was gut ist, welche Handlung ist dann jeweils richtig?

Es gibt eine gemeinsame Grundhaltung sowohl der klassischen-ethischen Tradition, wie sie von Platon und Aristoteles begründet wurde, als auch der ökonomischen Theorie praktischer Rationalität, die sich in der These zusammenfassen läßt: wenn die erste Frage beantwortet ist, wird die zweite Frage trivial. Diese triviale Lösung der zweiten Frage ist für die ökonomische Theorie nichts anderes als die Maximierung des erwarteten Nutzens. Die entgegengesetzte These, die in diesem Papier verteidigt werden soll, lautet: die zweite Frage ist mit der Beantwortung der ersten Frage nicht gelöst, ja die zweite Frage markiert einen eigenständigen Problemkomplex jeder normativen Handlungstheorie, der von der jeweils zugrundegelegten Theorie des Guten weitgehend unabhängig ist.

Um dies aufzuzeigen, ist es sinnvoll, den Begriff der kollektiven Rationalität einzuführen. Angenommen, es sei eine Relation "besser für eine Gruppe von Personen", eine kollektive Besser-Relation gegeben. Dann liegt es nahe, eine kollektive Handlung *a* (oder allgemeiner: eine kollektive Strategie) verstanden als Kombination individueller Handlungen kollektiv irrational zu nennen, wenn es eine andere kollektive Handlung *a'* gibt, die kollektiv besser ist als *a*. Für unsere Zwecke brauchen wir uns keine weiteren Gedanken um die Konstituenten der kollektiv-besser Relation machen; wir beschränken uns darauf, als Minimalbedingung ihre Pareto-Inklusivität zu fordern: eine kollektive Handlung, die nicht pareto-effizient ist, ist also kollektiv irrational (die Umkehrung gilt nicht). Nun gibt es einen Zweig der Entscheidungstheorie, die Logik kollektiver Entscheidung, der sich ausschließlich mit Adäquanzbedingungen an diese kollektivbesser Relation unter den Bedingungen einer individualistischen Methodologie befaßt. Trotz der für unsere Diskussion interessanten Ergebnisse dieses Forschungszweiges bleiben wir jedoch bei unserer asketischen Argumentationsstrategie: die Prämissen der Argumentation sollen so sparsam und damit so konsensual wie möglich sein.

Während sich der Begriff kollektiver Rationalität auf kollektive Handlungen bezieht, ist praktische Vernunft als ein Kriterium individueller Handlungen zu verstehen. Dennoch gibt es eine enge Verbindung zwischen kollektiver Rationalität und praktischer Vernunft: praktische Vernunft verlangt, daß sich eine individuelle Handlung in eine Struktur kollektiven Verhaltens einbetten läßt, die eine rationale Person akzeptieren kann. Es ergibt sich also erneut eine Inklusionsbeziehung: eine Kombination von *p* vernünftigen (d.h. praktischer

Vernunft gehorchenden) individuellen Handlungen ergibt eine kollektiv rationale kollektive Handlung (die Umkehrung gilt nicht). Ein Kernbereich praktischer Vernunft wird daher durch das Koordinationsproblem konstituiert. Individuelle Handlungen sind in einer Weise zu koordinieren, die pareto-effiziente Handlungen bedingen, dies ist eine notwendige, wenn auch keine hinreichende, Bedingung dafür, daß die individuellen Handlungen p-vernünftig sind.

Eine vollständig entwickelte Theorie praktischer Vernunft würde sicherlich nicht nur persönliche Interessen und subjektive Wertungen, sondern auch individuelle Rechte, Dispositionen und Tugenden der interagierenden Personen berücksichtigen. Persönlich meine ich, daß das Vertragsargument für die Entwicklung angemessener normativer Kriterien im Rahmen einer Theorie praktischer Vernunft eine wesentliche Rolle spielt, aber es ist nicht notwendig, diese kontraktualistische Theorie praktischer Vernunft hier näher zu charakterisieren. Im folgenden beschränken wir die Analyse vielmehr auf subjektive Präferenzen. Denn die Begrenztheit des ökonomischen Rationalitätskonzeptes wird gerade dann besonders deutlich, wenn man die Begrifflichkeit der Ökonomie beibehält. Unter dieser methodologischen Beschränkung gibt es eine enge Verbindung zwischen kollektiver Rationalität, die mit Hilfe des Begriffs der individuellen subjektiven Präferenz allein definiert ist und praktischer Vernunft: jedes System normativer Regeln hat die Bedingung zu erfüllen, vereinbar zu sein mit kollektiver Rationalität. Wenn ein System normativer Regeln (unter bestimmten Bedingungen) kollektive Handlungen nach sich zieht, die kollektiv irrational sind, so ist dieses System inadäquat. Daraus ergibt sich ein (negatives) Kriterium praktischer Vernunft: eine Handlung ist nicht p-vernünftig, wenn es kein System normativer Regeln gibt, welches diese Handlung erlaubt und das mit kollektiver Rationalität vereinbar ist.

### *3. Ökonomische Rationalität und praktische Vernunft in Entscheidungssituationen vom Typ des Gefangenendilemma*

In einigen Interaktionssituationen führt jedes individuell maximierende Verhalten der Beteiligten zu einem Ergebnis, das für alle schlechter ist, als wenn alle sich nicht individuell maximierend verhalten hätten. Von besonderem Interesse sind dabei solche Interaktionssituationen, wo dieses Dilemma nicht dadurch behoben werden kann, daß die Akteure mehr Information erhalten. Ein bekanntes Beispiel für eine Situation dieser Art ist das Gefangenendilemma. Die Geschichte, die diesem Typus von Interaktionssituation den Namen gegeben hat, muß hier nicht in Erinnerung gerufen werden. Wir wollen annehmen, daß beide Gefangenen zu vier Jahren Gefängnis verurteilt werden, wenn sie gestehen, zu je einem Jahr, wenn keiner gesteht, und, wenn nur einer gesteht, kommt dieser als Kronzeuge der Anklage frei, während der andere zu fünf Jahren verurteilt wird.

Im folgenden behalten wir die ursprüngliche Geschichte im Auge, da sie das Spannungsverhältnis ökonomischer Rationalität und praktischer Vernunft besonders dramatisch verdeutlicht. Ein Jahr Gefängnisarrest prägt die Lebenssituation einer Person derart einschneidend, daß anzunehmen ist, daß die Präferenzenerfüllung mit der zunehmenden Anzahl von Gefängnisjahren für beide Personen monoton fällt. Ökonomische Rationalität verlangt daher nicht kooperatives Verhalten (d.h. Gestehen). Ob ein Geständnis allerdings die in Abschnitt 1. aufgrund von Präferenzenkohärenz konstituierte Bewertungsfunktion maximiert, kann aufgrund dieser Information allein nicht entschieden werden. Es ist immerhin logisch möglich, daß ein Gefangener, der das Ziel hat, seinen Gefängnisarrest zu minimieren und dennoch kein Geständnis ablegt, Präferenzen zeigt, die alle in 1. genannten Axiome erfüllen. Insofern ist es logisch möglich, daß er die aufgrund der Kohärenzaxiome von 1. konstituierte Wertfunktion maximiert, ohne seine persönliche Präferenzenerfüllung zu optimieren.

Während die beiden Gefangenen ihre vier Jahre absitzen, werden sie sich vermutlich mit dem Vorwurf auseinandersetzen müssen, sie hätten sich unklug verhalten: es wäre für beide besser gewesen, wenn beide geschwiegen hätten. Während die klassische ökonomische Theorie praktischer Rationalität gezwungen war, ihr Verhalten als rational zu verteidigen, versuchen zeitgenössische Ökonomen, Wege zu ebnen, um diesen kontra-intuitiven Implikationen der ökonomischen Theorie zu entgehen. Um die Argumentation durchsichtig zu halten, ziehen wir im folgenden keine Superspiele in Betracht, sondern beschränken uns auf den einfachen Fall der einmaligen Entscheidung zweier Personen in einer Interaktionssituation vom Typ des Gefangenendilemmas.

Die je individuelle Optimierung der subjektiven Präferenzenerfüllung hat beiden Personen einen vierjährigen Gefängnisarrest eingebracht, obwohl die Alternative von jeweils einem Jahr Gefängnisarrest offengestanden hat. Unter der Annahme, daß die beiden Gefangenen sich rational verhalten, bieten sich nur zwei Möglichkeiten an, sie vor diesem Schicksal zu bewahren: entweder ist es nicht rational, seine Präferenzenerfüllung je individuell zu optimieren, oder (rationale) Kooperation in Situationen vom Typ des Gefangenendilemmas maximiert (je individuell) die Präferenzenerfüllung.

Das ökonomische Paradigma praktischer Rationalität läßt sich nur retten, wenn die zweite Möglichkeit besteht. Nun ist es eine Grundannahme der ökonomischen Theorie, daß eine Kritik von Entscheidungen immer zugleich eine Kritik (subjektiver) Präferenzen sei, da das Entscheidungsverhalten als eine Manifestation subjektiver Präferenzen verstanden wird. Daher wären diejenigen Präferenzen aufzudecken, die Kooperation und Maximierung kompatibel machen.<sup>12</sup> Und in der Tat wurde diese Suche nach den bis jetzt verborgenen Präferenzen unternommen - wir können diese Forschungsstrategie als Metapreferenz-Theorie bezeichnen. Wenn dieser Versuch erfolgreich wäre, würde sich das ökonomische Paradigma praktischer Rationalität auch seiner

ernsthaftesten Herausforderung gewachsen zeigen und damit seinen Anspruch bekräftigen, als generelle Konzeption praktischer Rationalität gelten zu können. Wenn dieser Versuch jedoch scheitert, ist keine andere Forschungsstrategie in Sicht, die das ökonomische Paradigma als generelle Konzeption praktischer Rationalität retten könnte.

Sen schlägt im Rahmen seiner Metapräferenztheorie zwei neue Präferenzordnungen vor, um die kooperative Lösung zu erreichen.<sup>13</sup> Das "Assurance Game" (AG): Mit "C" für kooperative Entscheidung und "D" für nichtkooperative ("defect") Entscheidung werden folgende Präferenzen der Personen A und B zugrunde gelegt.

$$A: (C,C) > (D,C) > (D,D) > (C,D)$$

$$B: (C,C) > (C,D) > (D,D) > (D,C)$$

Jede Person zieht es also vor, sich kooperativ zu verhalten, vorausgesetzt, die andere Person verhält sich ebenfalls kooperativ. Man könnte auch sagen, die Erwartung, daß sich die andere Person kooperativ verhält, ist eine notwendige Bedingung dafür, daß man sich selbst kooperativ verhält (daher die Bezeichnung "assurance game").

Zum Vergleich die Präferenzen des ursprünglichen Gefangenendilemmas (PD):

$$A: (D,C) > (C,C) > (D,D) > (C,D)$$

$$B: (C,D) > (C,C) > (D,D) > (D,C)$$

Während das ursprüngliche Gefangenendilemma nur einen Gleichgewichtspunkt in (D,D) hatte, hat AG zwei Gleichgewichtspunkte, (C,C) und (D,D). Da die individuellen Entscheidungen jedoch davon abhängen, welche Entscheidung man von der anderen Person erwartet, ist auch in AG nicht garantiert, daß die kooperative Lösung erreicht wird, obwohl die kooperative Lösung (C,C) pareto-besser ist als die Kombination nicht-kooperativen Verhaltens (D,D).

Unter der Bedingung, daß beide Personen die gleichen subjektiven Erwartungen bezüglich der Entscheidung des anderen haben, ergeben sich in AG zwei Gleichgewichtspunkte. Es ist ein wesentliches Charakteristikum dieses Spieles, daß das kooperative Ergebnis durch eine Übereinkunft der Teilnehmer erreicht werden kann, die ohne äußere Sanktionen auskommt (im Gegensatz zu PD). Nach wie vor von einer Gefangenendilemmasituation ausgehend, betrachtet Sen eine weitere Präferenzenkonstellation, die das sogenannte "Other Regarding-Game" (OR) konstituiert:

$$A: (C,C) > (C,D) > (D,C) > (D,D)$$

$$B: (C,C) > (D,C) > (C,D) > (D,D)$$

Für OR gilt nun, daß unabhängig davon, welche Erwartungen die Personen bezüglich des Entscheidungsverhaltens der anderen haben, jede die kooperative Entscheidung wählt. Das kooperative Ergebnis ist in diesem Fall dominant, während in der ursprünglichen Gefangenendilemmasituation das nicht-kooperative Ergebnis dominant war. Wenn die wechselseitige Erwartung besteht, die andere Person verhalte sich kooperativ, dann ist (C,C) nicht nur rational in Bezug auf die AG-Präferenz, sondern auch in Bezug auf die ursprünglichen PD-Präferenzen. Wenn sich also die beiden Personen so verhalten würden, als ob sie AG-Präferenzen hätten, so käme es zu einem kollektiv-rationalen Ergebnis bezüglich des ursprünglichen Gefangenendilemmas während sich ja ein kollektiv irrationales Ergebnis einstellt, wenn sie ihre Entscheidung an den ursprünglichen Präferenzen des Gefangenendilemmas orientieren.

Im Altruissusspiel hingegen ist die Bedingung des wechselseitigen Vertrauens in die Kooperationsbereitschaft entbehrlich. Hier ist die kooperative kollektive Entscheidung dominant. Diese Entscheidung ist darüberhinaus kollektiv-rational, und zwar sowohl bezüglich der individuellen Präferenzen des Altruissusspiels als auch bezüglich der AG-Präferenzen.

Sowohl bezüglich der AG- wie bezüglich der OR-Präferenzen ist das kooperative Ergebnis das je individuell günstigste. Für die PD Präferenzen gilt das nur unter der Bedingung der Symmetrie: es ist bezüglich PD individuell besser, wenn beide sich kooperativ, als wenn beide sich unkooperativ verhalten. AG-Präferenzen garantieren kollektive Rationalität bezüglich der (ursprünglichen) PD Präferenzen unter der epistemischen Bedingung wechselseitiger Kooperationserwartung, während OR-Präferenzen kollektive Rationalität ohne jede epistemische Bedingung garantieren.

Offensichtlich gibt es also Interaktionssituationen, in denen ökonomisch-rationales Verhalten bezüglich anderer Präferenzen als sie die handelnden Personen tatsächlich haben, zu kollektiver Rationalität bezüglich dieser tatsächlichen Präferenzen führen. Ist es deshalb sinnvoll, eine moralische Rangordnung unterschiedlicher Präferenzen vorzunehmen, je nachdem, ob sie kollektive Rationalität garantieren, ermöglichen oder ausschließen, wie Sen vorschlägt?

Moralische Metapräferenzen können auf dreierlei Weise interpretiert werden, die sorgfältig voneinander unterschieden werden müssen.

1. Man kann moralische Metapräferenzen als Empfehlung interpretieren, sich bestimmte Präferenzen anzueigenen, nach denen sich dann das Verhalten (in Anwendung entscheidungs- und spieltheoretischer Kriterien) richten sollte.
2. Moralische Metapräferenzen können interpretiert werden als Vorschlag, sich so zu verhalten, als ob man bestimmte Präferenzen hätte, auch wenn man tatsächlich andere Präferenzen hat.

3. Moralische Metapräferenzen können jedoch auch als Relation zwischen Präferenzen unterschiedlichen Typs verstanden werden.

Jede dieser Interpretationen birgt spezifische Schwierigkeiten.

**ad 1:** Dieser Versuch, kollektive Rationalität qua Maximierung neuer Präferenztypen sicherzustellen, scheitert von vornherein, da die ursprüngliche Interaktionssituation des Gefangenendilemmas bei dieser Interpretation restlos verschwindet.

**ad 2:** Auf der Grundlage der tatsächlichen Präferenzen könnte moralisches Verhalten als ökonomisch-rationales Verhalten bezüglich bestimmter fiktiver Präferenzrelationen konzipiert werden. Moralisches Verhalten wäre dann nicht ökonomisch rational bezüglich der tatsächlichen Präferenzen, aber ökonomisch rational bezüglich dieser fiktiven Präferenzen und würde damit bezüglich der tatsächlichen Präferenzen kollektive Rationalität ermöglichen oder sogar sicherstellen. Dementsprechend gibt es geeignetere und weniger geeignete fiktive Präferenzrelationen, die von einer Metapräferenzrelation in eine Rangordnung gebracht werden.

Die fiktiven Präferenzen werden danach beurteilt, in welchem Grade sie das moralisch wünschenswerte Ergebnis für den Fall sicherstellen, daß sich die Akteure bezüglich dieser fiktiven Präferenzen maximierend verhalten. Daraus ergibt sich natürlich keineswegs, daß diese fiktiven Präferenzen eine Rangordnung von Zuständen (hier: isomorph mit kollektiven Handlungen) bilden, die einer (unmittelbaren) moralischen Beurteilung dieser Zustände entspricht. In diesem Falle gäbe es keinen vernünftigen Grund, sich diese Präferenzen zu eigen zu machen, außer dem, daß sie, wenn man maximierendes Verhalten bezüglich dieser Präferenzen zugrundelegt, zu einem kollektiv rationalen Ergebnis führen. Dann aber verliert dieser Umweg über fiktive Präferenzen jede Plausibilität, denn dann wäre die Empfehlung, die eigene (individuelle) Entscheidung an kollektiver Rationalität auszurichten, ausreichend. Das Modell der fiktiven Präferenzen ist bei dieser Interpretation ausschließlich damit zu motivieren, die Auffassung zu retten, rationales Verhalten sei grundsätzlich maximierendes Verhalten. Da es für das Kriterium praktischer Rationalität keinen Beitrag leistet und nicht anzunehmen ist, daß es in irgendeinem Zusammenhang mit der tatsächlichen Motivationsstruktur steht, dient es hier offensichtlich ausschließlich der Immunisierung der maximierenden Auffassung praktischer Rationalität.

**ad 3:** Diese Interpretation beruht auf der Annahme, daß es unterschiedliche Arten von Präferenzen gibt, und daß man Einzelpersonen simultan voneinander abweichende Präferenzen dieser unterschiedlichen Typen zuordnen kann. So können Personen in der Regel zwischen ihren persönlichen Interessen und ihren subjektiven moralischen Vorstellungen unterscheiden. Demnach wäre zumindest

zwischen Präferenzen, die die eigenen Interessen repräsentieren, und Präferenzen, die die moralischen Beurteilungen repräsentieren, zu unterscheiden. Bei der Beurteilung eines Zustandes kann es aber auch Abwägungen geben zwischen persönlichen Interessen und moralischen Aspekten. Das Ergebnis dieser Abwägung sei im folgenden als "**resultierende Präferenz**" bezeichnet. Interaktionssituationen wären demnach im allgemeinen von (mindestens) drei Präferenz-Strukturen charakterisiert: eine Interessenstruktur, eine moralische Bewertungsstruktur und eine Struktur resultierender Präferenzen. Die Gefangenendilemma-Situation sei im folgenden durch eine entsprechende Interessenstruktur charakterisiert. Ob eine Gefangenendilemma-Situation vorliegt, ist bei dieser Interpretation nicht mehr allein aufgrund des gezeigten Entscheidungsverhaltens bestimmbar, denn es ist ja immerhin möglich, daß das Entscheidungsverhalten einer rationalen Person von ihren persönlichen Interessen nicht allein bestimmt ist. Wenn etwa die subjektive Moral der beiden Gefangenen pareto-inklusiv ist, dann ist für sie der Zustand, der sich aus beiderseitig kooperativem Verhalten ergibt moralisch besser, als der Zustand, der sich beim beiderseitigen unkooperativen Verhalten einstellt. Eine Abwägung dieser beiden Aspekte führt zu entsprechenden resultierenden Präferenzen. Eine der vorliegenden Präferenz-Strukturen muß jedoch über das Maximierungskriterium mit dem individuellen Entscheidungsverhalten verkoppelt bleiben, wenn mit der Einführung von Metapräferenzen das ökonomische Modell praktischer Rationalität gerettet werden soll. Die naheliegende Interpretation wäre demnach: Metapräferenzen empfehlen auf der Grundlage einer vorliegenden Interessenstruktur bestimmte resultierende Präferenz-Strukturen, wobei das maximierende Konzept praktischer Rationalität bezüglich der resultierenden Präferenzen beibehalten wird. Eine andere Möglichkeit wäre, die moralischen Präferenzen handlungsleitend zu machen und damit an einem dichotomischen Konzept festzuhalten. Diese Möglichkeit ist jedoch mit moralpsychologischen Beobachtungen nur schwer in Einklang zu bringen.

Erst mit der Differenzierung des Präferenzbegriffes erhält der Ansatz der Metapräferenzen eine plausible Interpretation. Die sorgfältige Durchführung dieses Ansatzes zeigt jedoch, daß damit das ökonomische Modell praktischer Rationalität nicht gerettet werden kann, ja es zeigt meines Erachtens sogar recht deutlich, woran das liegt: im Rahmen des Maximierungsmodells läßt sich kollektive Rationalität nicht auf individuelle Rationalität reduzieren.

Jede Person hat im Gefangenendilemma 4!, also 24 Möglichkeiten, die vier möglichen Ergebnisse des Spiels in eine Präferenzordnung zu setzen. Insgesamt ergeben sich also  $24 \times 24 = 576$  logisch mögliche Kombinationen individueller Präferenzordnungen. Bei drei Präferenz-Strukturen hat dies  $576^3 (= 191102976)$  logisch mögliche Kombinationen dieser Präferenz-Strukturen zur Folge. Es ist anzunehmen, daß ein Großteil dieser Präferenz Strukturen aufgrund psychologischer Sachverhalte empirisch nicht vorkommt. Wenn wir jedoch die

Interessenstruktur als gegeben voraussetzen und einen weiteren Präferenz-Typ als handlungsleitend annehmen, bleiben immer noch 576 Möglichkeiten. Um das Scheitern dieses Ansatzes aufzuzeigen, genügt es jedoch, sich auf diejenigen Fälle zu konzentrieren, die von Symmetrie gekennzeichnet sind. Wir legen also jeweils eine bestimmte Relation zwischen Interessenstrukturen und resultierenden individuellen Präferenzen für beide Beteiligten gleichermaßen zugrunde.

Eine naheliegende Rationalitätsbedingung für diese interpräferenzuelle Relation wäre die Pareto-Inklusivität bezüglich der zugrundeliegenden Interessen-Strukturen. Eine Person würde demnach einen Zustand  $x$ , einem Zustand  $y$  resultierend vorziehen, wenn alle an der Interaktion Beteiligten ein (individuelles) Interesse daran haben, daß  $x$  und nicht  $y$  realisiert wird. Wir nennen dieses Kriterium "PRW" (Pareto-Inklusivität der resultierenden Präferenzen bezüglich der (individuellen) Wohlfahrtspräferenzen).

1. Die Personen richten ihre resultierenden Präferenzen ausschließlich an ihren persönlichen Interessen aus. Man könnte diese resultierenden Präferenzen egoistische Präferenzen nennen.

$$\begin{aligned} \text{A: } & (D,C) > (C,C) > (D,D) > (C,D) \\ \text{B: } & (C,D) > (C,C) > (D,D) > (D,C) \end{aligned}$$

Diese Präferenzen erfüllen PRW. Tragik der Egoisten (und Inhalt des Gefangenen-Dilemmas) ist es, daß sie mit diesen Präferenzen bei (D,D) enden und daher ihren eigenen Interessen gerade dadurch schaden, daß sie diese zum einzigen Kriterium ihrer Entscheidungen machen.

2. Die resultierenden Präferenzen sind von einem starken Gerechtigkeitsempfinden geprägt.

$$\begin{aligned} \text{A: } & (C,C) > (D,D) > (D,C) > (C,D) \\ \text{B: } & (C,C) > (D,D) > (C,D) > (D,C) \end{aligned}$$

Hier wird angenommen, daß die Personen, falls sie unter moralischen Gesichtspunkten keinen Unterschied zwischen zwei Zuständen machen können, den Zustand resultierend vorziehen, in dem es ihnen persönlich besser geht. Diese resultierenden Präferenzen garantieren nicht kollektive Rationalität. Denn jede Person verhält sich unkooperativ, wenn sie sich vom anderen kooperatives Verhalten erwartet. Sowohl (C,C), wie (D,D) sind Gleichgewichtspunkte dieses Spiels.

3. Auch, wenn sich der Gerechtigkeitsinn mit Altruismus verbindet und sich folgende folgende resultierende Präferenzen ergeben:

$$\begin{aligned} \text{A: } & (C,C) > (D,D) > (C,D) > (D,C) \\ \text{B: } & (C,C) > (D,D) > (D,C) > (C,D) \end{aligned}$$

ist die kollektiv rationale Entscheidung aus den gleichen Gründen wie unter (2) nicht garantiert.

4. Den Gegenpol zu diesen an Gerechtigkeit orientierten Präferenzen bilden solche, die man als Ausdruck eines stark ausgeprägten Anti-Egalitarismus interpretieren könnte.

$$\begin{aligned} \text{A: } & (D,C) > (C,D) > (C,C) > (D,D) \\ \text{B: } & (C,D) > (D,C) > (C,C) > (D,D) \end{aligned}$$

Diese anti-egalitären resultierenden Präferenzen haben zwei Gleichgewichtspunkte (C,D) und (D,C).

5. Verdrängt jedoch der Altruismus den Gerechtigkeitssinn in der Gestalt folgender resultierender Präferenzen:

$$\begin{aligned} \text{A: } & (C,C) > (C,D) > (D,C) > (D,D) \\ \text{B: } & (C,C) > (D,C) > (C,D) > (D,D) \end{aligned}$$

so ist endlich kollektive Rationalität garantiert: Die kollektiv rationale Entscheidung ist Gleichgewichtspunkt in dominanten Strategien. Allerdings kommen einem gewisse Zweifel an der inneren Konsistenz dieser Präferenzrelation, denn  $(D,C) > (D,D)$  ist weder altruistisch noch gerecht, sondern egoistisch, während  $(C,D) > (D,C)$  extrem altruistisch ist.

6. Aber auch ein konsequenter Altruismus garantiert kollektive Rationalität in Interaktionssituationen von der Art des Gefangenendilemmas:

$$\begin{aligned} \text{A: } & (C,D) > (C,C) > (D,D) > (D,C) \\ \text{B: } & (D,C) > (C,C) > (D,D) > (C,D) \end{aligned}$$

7.

$$\begin{aligned} \text{A: } & (C,C) > (C,D) > (D,D) > (D,C) \\ \text{B: } & (C,C) > (D,C) > (D,D) > (C,D) \end{aligned}$$

Diese symmetrische Präferenz-Struktur könnte bei kardinaler Interpretation der Interessen-Struktur als Ausdruck eines abgeschwächten Altruismus interpretiert werden: Wenn die Erhöhung des Wohlfahrtsniveaus der anderen Person gleich hoch wie die Absenkung des eigenen Wohlfahrtsniveaus ist, wird die kollektive Entscheidung zugunsten des anderen vorgezogen. Ist die Veränderung des

eigenen Wohlfahrtniveaus jedoch doppelt so groß wie die des anderen, neigen sich die resultierenden Präferenzen zu den eigenen Gunsten. Diese Präferenzen haben einen Gleichgewichtspunkt in (C,C) und garantieren also immer noch kollektive Rationalität.

8. Die resultierenden Präferenzen, die dem Assurance-Game (AG) entsprechen, haben folgende Struktur:

$$\begin{aligned} \text{A: } & (C,C) > (D,C) > (D,D) > (C,D) \\ \text{B: } & (C,C) > (C,D) > (D,D) > (D,C) \end{aligned}$$

Diese Personen sind bereit, sich kooperativ zu verhalten, wenn sie annehmen, daß sich die andere Person ebenfalls kooperativ verhält. Damit wird die kollektiv rationale Entscheidung (C,C) ein Gleichgewichtspunkt der Interaktionssituation. Allerdings ist die kollektiv irrationale Entscheidung (D,D) ebenfalls ein Gleichgewichtspunkt. Kein Beteiligter hat eine dominante individuelle Strategie. Hier ist wechselseitiges Vertrauen Bedingung kollektiver Rationalität.

Die restlichen vier Präferenz-Strukturen sind schwieriger zu interpretieren:

9.

$$\begin{aligned} \text{A: } & (C,C) > (D,C) > (C,D) > (D,D) \\ \text{B: } & (C,C) > (C,D) > (D,C) > (D,D) \end{aligned}$$

Diese Präferenz-Struktur könnte man als ordinale Repräsentation utilitaristischer Präferenzen interpretieren - sie garantieren kollektive Rationalität: CC ist Gleichgewichtspunkt in dominanten Strategien.

10.

$$\begin{aligned} \text{A: } & (C,D) > (D,C) > (C,C) > (D,D) \\ \text{B: } & (D,C) > (C,D) > (C,C) > (D,D) \end{aligned}$$

11.

$$\begin{aligned} \text{A: } & (D,C) > (C,C) > (C,D) > (D,D) \\ \text{B: } & (C,D) > (C,C) > (D,C) > (D,D) \end{aligned}$$

Diese beiden Präferenz-Strukturen ergeben wiederum zwei Gleichgewichtspunkte: (C,D) und (D,C).

Die letzte PRW-rationale Präferenz-Struktur garantiert erneut kollektive Rationalität:

12.

A: (C,D) &gt; (C,C) &gt; (D,C) &gt; (D,D)

B: (D,C) &gt; (C,C) &gt; (C,D) &gt; (D,D)

Betrachten wir die PRW-irrationalen (symmetrischen) Präferenz-Strukturen.

13. Zunächst Präferenzen, die den Wunsch, sich selbst zu schaden, offenbaren:

A: (C,D) &gt; (D,D) &gt; (C,C) &gt; (D,C)

B: (D,C) &gt; (D,D) &gt; (C,C) &gt; (C,D)

Selbsthaß garantiert zusammen mit reinem (6) und abgeschwächtem (7) Altruismus, den etwas seltsam motivierten Präferenzen OR (5) und zwei weitere recht unplausible Präferenz-Strukturen (9) und (12), die kollektiv rationale Lösung. Kollektiv rational bezieht sich natürlich immer auf die Interessen-Struktur der Situation. Im Falle masochistischer Präferenzen heißt dies, daß die Personen mit dem Ergebnis nicht zufrieden sein können, denn, obwohl sie den Zustand, in dem es ihnen persönlich am schlechtesten geht, an die Spitze, und den Zustand, in dem es ihnen am besten geht, an das Ende ihrer (resultierenden) Präferenzen gesetzt haben, wird (C,C) realisiert. Das Gefangenendilemma reproduziert sich gewissermaßen für Masochisten auf der Ebene der resultierenden Präferenzen.

14. Das Gegenstück sadistischer Präferenzen

A: (D,C) &gt; (D,D) &gt; (C,C) &gt; (C,D)

B: (C,D) &gt; (D,D) &gt; (C,C) &gt; (D,C)

verstößt wie Masochismus gegen die Bedingung PRW. Während jedoch Egoismus im Gefangenendilemma nicht erfolgreich war, ist das wechselseitige Bestreben, dem Anderen zu schaden, so weit erfolgreich, wie es die Symmetrie der Situation zuläßt: Nicht-Kooperation ist dominant. Dennoch ist der Gleichgewichtspunkt (D,D) bezüglich der Interessen-Struktur kollektiv irrational. Nun ist man versucht, in diesem Fall die kollektive Irrationalität der nicht kooperativen kollektiven Entscheidung zu bezweifeln. Dabei würde man jedoch die Tatsache übersehen, daß die nicht kooperative Entscheidung auch für den "Sadisten" schlechter ist als die kooperative, denn andernfalls wären die Wohlfahrtsbewertungen abzuändern und wir hätten es nicht mehr mit einem Gefangenendilemma bezüglich der Interessen-Struktur der Situation zu tun.

Immerhin findet sich unter den 12 PRW-irrationalen Präferenz-Strukturen eine Struktur, die kollektive Rationalität garantiert (13). Dies ist auffallend, da ja PRW-Irrationalität bedeutet, daß die Gefangenen (D,D) dem pareto-optimalen

(C,C) vorziehen. Immerhin sechs PRW-irrationale Strukturen erlauben kollektive Rationalität:

15.

A: (D,D) > (C,C) > (D,C) > (C,D)

B: (D,D) > (C,C) > (C,D) > (D,C)

16.

A: (D,D) > (C,C) > (C,D) > (D,C)

B: (D,D) > (C,C) > (D,C) > (C,D)

17.

A: (D,D) > (C,D) > (C,C) > (D,C)

B: (D,D) > (D,C) > (C,C) > (C,D)

18.

A: (D,C) > (C,D) > (D,D) > (C,C)

B: (D,D) > (D,C) > (D,C) > (C,C)

19.

A: (C,D) > (D,C) > (D,D) > (C,C)

B: (D,C) > (C,D) > (D,D) > (C,C)

20.

A: (C,D) > (D,D) > (D,C) > (C,C)

B: (D,C) > (D,D) > (C,D) > (C,C)

Fünf der PRW-irrationalen Strukturen führen unweigerlich zu kollektiver Irrationalität. Neben den sadistischen Präferenzen (14) sind dies die folgenden:

21.

A: (D,D) > (D,C) > (C,C) > (C,D)

B: (D,D) > (C,D) > (C,C) > (D,C)

22.

A: (D,C) > (D,D) > (C,D) > (C,C)

B: (C,D) > (D,D) > (D,C) > (C,C)

23.

A: (D,D) > (D,C) > (C,D) > (C,C)

B: (D,D) > (C,D) > (D,C) > (C,C)

24.

A: (D,D) > (C,D) > (D,C) > (C,C)

B: (D,D) > (D,C) > (C,D) > (C,C)

Die ursprüngliche Hoffnung kollektive Rationalität über Präferenzen sicherzustellen, die moralischen Maßstäben eher genügen als egoistischen, hat sich also nicht erfüllt.

Diese 24 (symmetrischen) Präferenz-Strukturen zerfallen in drei Klassen, die sich darin unterscheiden, ob kollektive Rationalität als Ergebnis maximierender Entscheidungen garantiert, möglich oder ausgeschlossen ist. In jeder dieser drei Klassen befinden sich PRW-rationale und PRW-irrationale Präferenzrelationen. In der ersten Klasse der kollektive Rationalität garantierenden Präferenz-Strukturen befinden sich nicht nur die ehrenwerten altruistischen Präferenz-Strukturen (6) und (7), sondern auch die PRW-irrationalen masochistischen Präferenzen (13) und eine ganze Gruppe wenig plausibler Strukturen, etwa (9), (11) und (12). In der zweiten Klasse der kollektive Rationalität ermöglichenden Präferenz-Strukturen finden sich solche wie (2) und (3), die intuitiv als moralisch höherrangig empfunden werden als einige der Präferenz-Strukturen der ersten Gruppe - etwa (13) -, aber auch solche, die den eigenen Interessen Vorrang einräumen (8) - und schließlich eine große Gruppe von PRW-irrationalen Präferenz-Strukturen (15 bis 20). In der dritten Klasse der kollektive Rationalität ausschließenden (kein Gleichgewichtspunkt in kooperativen Strategien) Präferenz-Strukturen rangieren egoistische und anti-egalitäre, aber auch die PRW-irrationalen sadistischen und schließlich vier PRW-irrationale Präferenzstrukturen - (21), (22), (23), (24).

Der Ansatz der Metapräferenzen verliert seine Plausibilität, sobald man die Betrachtung nicht nur auf wenige, ausgewählte Präferenz-Strukturen beschränkt, sondern konsequent alle Präferenz-Strukturen danach beurteilt, in welchem Grad sie, ökonomische Rationalität vorausgesetzt, zu kollektiver Rationalität beitragen. Es wird dann deutlich, daß dieser Umweg, kollektive Rationalität zu sichern, nicht überzeugen kann. Der Konflikt zwischen individueller und kollektiver Rationalität läßt sich in adäquater Weise nicht durch eine ergebnisorientierte Veränderung subjektiver Präferenzen, die das Verhalten qua Maximierung steuern, lösen. Aufgrund der engen Verbindung zwischen praktischer Vernunft und kollektiver Rationalität läßt sich daher p-vernünftiges Verhalten nicht als Maximierung einer adäquaten subjektiven Wertfunktion rekonstruieren.

Dennoch scheint es nicht ausgeschlossen, daß die unter (1) aufgeführten Kohärenzaxiome auch von p-vernünftigem Verhalten erfüllt werden. Damit wäre gesichert, daß es eine Präferenzrelation gibt, die von einer quantitativen Bewertungsfunktion repräsentiert werden kann, welche durch p-vernünftiges Verhalten maximiert wird. Es liegt aufgrund der oben analysierten Präferenz-Strukturen auf der Hand, daß diese Präferenzrelationen dem betreffenden p-vernünftigen Akteur sinnvollerweise nicht zugeschrieben werden können. Denn das Handlungskriterium des p-vernünftigen Akteurs orientiert sich in dem hier betrachteten Kernbereich an kollektiver Rationalität. Ein solcher Akteur wird sich die formal zuschreibbaren Präferenzen nicht zu eigen machen. Da zur

Bestimmung kollektiver Rationalität die jeweilige Interaktionsstruktur ausschlaggebend ist, ließen sich solche formal zuschreibbaren Präferenzen nicht auf Parameter reduzieren, wie sie für die ökonomische Theorie der Rationalität paradigmatisch sind. Der p-vernünftige Akteur orientiert seine Handlung in einer Interaktionssituation vom Typ des Gefangenen-Dilemmas an kollektiver Rationalität. In der Regel wird dies seine Präferenzen bezüglich der vier möglichen Ergebnisse des 'Spiels' nicht verändern. Dies ist auch nicht notwendig, denn die p-Vernünftigkeit einer Entscheidung reicht als Handlungsmotiv aus.

*Anmerkungen*

- 1 Vgl. Aristoteles, Nikomachische Ethik 1146 ff., dazu Verf., "Der praktische Schluß des Aristoteles und das revealed preference Konzept der rationalen Entscheidungstheorie", in: Praktische Logik, hg. von P. Klein, S.203-216.
- 2 Vgl. J. Bentham, An Introduction to the Principles of Morals and Legislation, London 1789, Kap. I-V; a. (posthum), Deontology, hg. von J. Bouwring, London, Edinburgh 1834; J.S. Mill, Utilitarianism, London 1863, Kap. IV.
- 3 Vgl. etwa R.B. Brandt, Ethical Theory, Englewood Cliffs, N.J. 1959; J. Rawls, "Two Concepts of Rules", Philosophical Review 64 (1955) 3-32; N. Hoerster, Utilitaristische Ethik und Verallgemeinerung, Freiburg/ München 1977.
- 4 Dazu kritisch A.K. Sen, "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory", in: Scientific Models and Man: The Herbert Spencer Lectures 1976, hg. von H. Harris, Oxford 1978.
- 5 Vgl. J. Elster, Ulysses and the Sirens, Cambridge 1979 und Sour Grapes, Cambridge 1983, in den hier relevanten Teilen ins Deutsche übertragen in: Subversion der Rationalität, Frankfurt/ New York 1987, sowie Solomonic Judgements, Cambridge 1989.
- 6 Vgl. A.K. Sen, Collective Choice and Social Welfare, San Francisco u.a. 1970; a. K. Suzumura, Rational Choice, Collective Decisions, and Social Welfare, Cambridge 1983.
- 7 Vgl. B. Williams/ J.J.C. Smart, Utilitarianism - For and Against, Cambridge 1973; S. Scheffler, Consequentialism and Its Critics, Oxford 1988; Verf., Der Konsequentialismus - Rekonstruktion und Kritik, Habil. (1989) i. E.
- 8 Vgl. A. MacIntyre, Der Verlust der Tugend, Frankfurt /M. 1987; R.M. Hare, Moral Thinking. Its Levels, Method and Point, Oxford 1981.
- 9 Vgl. etwa M. Allais/ O. Hagen (Hg.), Expected Utility Hypothesis and the Allais Paradox, Dordrecht 1979.
- 10 Vgl. J. von Neumann/ O. Morgenstern, Theory of Games and Economic Behavior, Princeton 1947.
- 11 M. Olson, The Logic of Collective Action, Cambridge 1965.

- 12 Vgl. W. Raub/ T. Voss, Conditions for Cooperation in Problematic Situations, in: Paradoxical Effects of Human Behavior, hg. von A. Diekmann u. P. Mitter, Heidelberg 1986; sowie: T. Raub, Rationale Akteure und soziale Institutionen, München 1985.
- 13 A.K. Sen, Choice, Orderings and Morality, in: Practical Reason, hg. von S. Körner, Oxford 1974; vgl. a. R.C. Jeffrey, Preferences among Preferences, Journal of Philosophy 71 (1984) 377-391; T.C. Schelling, Egonomics, American Economic Review (Papers & Proceedings) 68 (1978) 290-294; ders. The Intimate Contest of Self-Command, Public Interest 60 (1980); ders. Choice and Consequence, Cambridge/Mass. 1984; D. George, Metapreferences: Reconsidering Contemporary Notions of Free Choice, Journal of Social Economics 11 (1984) 92-107.

# Namensregister

- Albert, H. 13, 20, 22  
Aldrup, D. 12, 20, 22  
Allais, M. 151, 207  
Anand, P. 113  
Anscombe, F.J. 90, 99, 100  
Apel, K.-O. 20, 22  
Aristoteles 17, 131, 138, 151  
Arrow, K. 8, 37, 40, 63, 73, 84, 89, 100, 112-114  
Atkinson, A.B. 84  
Aumann, R. 90, 99, 100  
Axelrod, R. 17, 18, 21-23, 54, 60, 133, 159, 171, 172, 185, 189
- Baigent, N. 6, 113, 205, 217  
Becker, G.S. 14, 21, 22  
Bellman, R. 39, 40  
Bentham, J. 151, 201, 206  
Blackwell, D. 100  
Bouwring, J. 151  
Brandt, R.B. 151  
Broad, C.D. 170, 172  
Buchanan, J.M. 17, 21, 22, 24, 39, 40, 133  
Burke 115, 127  
Butler, J. 134
- Carnap, R. 88, 100  
Chatterjee, S. 113
- De Finetti, B. 85, 88, 100  
Deb, R. 72, 73  
Descartes, R. 13  
Diekmann, A. 152, 188-190  
Dilthey, W. 7, 115  
Dworkin, R. 175, 181
- Elster, J. 17, 21, 22, 84, 113, 114, 151
- Feinberg, J. 72, 73

- Friedman, J.W. 189  
 Fuller, L.L. 181
- Gaertner, W. 5, 6, 9, 66, 67, 71, 73, 113, 114, 217  
 Gähde, U. 171, 172  
 Gärdenfors, P. 71, 73, 207  
 Gauthier, D. 3, 17, 18, 21, 22, 33, 37, 39, 40, 44- 52, 54, 55, 58-60, 113, 114  
 Gehlen, A. 17, 21, 22  
 George, D. 152  
 Gibbard, A. 6, 63, 64, 67, 68, 72, 73  
 Girshick, M.A. 100  
 Good, I.J. 88, 100  
 Gowans, C.W. 171, 172  
 Griffin, J. 200, 207
- Habermas, J. 20-22, 181  
 Hagen, O. 151, 207  
 Hardin, R. 188, 189  
 Hare, R.M. 132, 151  
 Harris, H. 151  
 Harsanyi, J.G. 44, 60, 113, 114  
 Hayek, F.A. von 18  
 Hegselmann, R. 8, 113, 114, 189, 190, 217  
 Heiner, R.A. 17, 21, 22  
 Helvetius, C.A. 115  
 Hobbes, T. 17, 43, 44, 60, 61, 118, 120, 206  
 Hoerster, N. 151  
 Hogarth, M. 113, 114  
 Hollis, M. 7, 130, 156, 170-172, 217  
 Homann, K. 5, 20-22, 181, 217  
 Hume, D. 120, 121, 123, 129, 192, 206
- Jahn, J. 113, 114  
 Jaynes, E.T. 88, 100  
 Jeffrey, R.C. 152  
 Jeffreys, H. 88, 100
- Kadane, J.B. 90-94, 96-98, 100  
 Kalai, E. 33, 39, 40, 49, 61  
 Kant, I. 8, 20, 22, 55, 114, 115, 121, 128, 131, 155, 156, 172, 183, 184, 190  
 Kliemt, H. 17, 21, 22, 188, 190

- Koertge, N. 13, 20, 23  
Kohlberg, L. 18, 19, 21, 23  
Kolm, S. 113, 114  
Körner, S. 152  
Krabs, W. 113, 114  
Krause, U. 7, 113, 114, 217  
Kuhn, T.S. 13, 14, 20, 21, 23  
Kutschera, F.v. 189, 190
- Lakatos, I. 13, 20, 23  
Levi, I. 6, 7, 85, 89, 100, 103, 113, 114, 172, 217  
Locke, J. 43, 45, 60, 61, 117, 207  
Luce, R.D. 39, 40
- Machiavelli, N. 119, 125  
MacIntyre, A. 129, 130, 151  
Mackie, J.L. 188, 190  
Marwell, G. 40  
Maynard Smith, J. 53, 61  
McClenen, E.F. 5, 39, 40, 207, 217  
McLean, I. 171, 172  
Meckling, W.H. 20, 23  
Mill, J.S. 63-65, 73, 151, 199, 206, 207  
Mitter, P. 152, 188-190  
Morgenstern, O. 39, 41, 48, 61, 133, 151
- Nash, J.F. 33, 39-41, 45, 49, 50, 61  
Neumann, J. von 39, 41, 48, 61, 151  
Nida-Rümelin, J. 7, 217  
North, D.C. 16, 21, 23  
Nozick, R. 71, 73, 191, 207
- Olson, M. 21, 23, 32, 41, 133, 151
- Pähler, K. 13, 20, 21, 23  
Pattanaik, P.K. 66, 67, 71-73  
Platon 1, 17, 18, 138
- Ramsey, F.P. 88, 100  
Rapaport, A. 185, 188-190  
Raub, W. 21-23, 152, 188, 189, 190  
Rawls, J. 3, 33-35, 37, 39, 41, 44, 58, 59, 61, 151, 191, 194-197, 207  
Raz, J. 180, 181

- Reder, M.W. 25, 37, 41, 113, 114  
 Rescher, N. 13, 20, 23  
 Rousseau, J.J. 126-128
- Sartre, J.P. 88, 100  
 Satterthwaite, M.A. 63, 73  
 Savage, L.J. 88, 100  
 Sayre McCord, G. 40, 41  
 Scanlon, T. 193, 194, 206, 207  
 Scheffler, S. 151  
 Schelling, T.C. 17, 21, 23, 56, 61, 75, 84, 113, 114, 152  
 Schervish, M.J. 90-94, 96-98, 100  
 Schick, F. 114  
 Schluchter, W. 15, 21, 23  
 Schnädelbach, H. 20, 22, 23  
 Schumpeter, J.A. 12, 15, 16, 20, 21, 23  
 Seidenfeld, T. 90-94, 96-98, 100  
 Seidl, C. 21  
 Sen, A.K. 5, 6, 63-65, 67-73, 75, 81, 83, 84, 113, 114, 141, 142, 151, 152,  
 188-190, 206, 207  
 Seneta, E. 113  
 Sidgwick, H. 155, 171, 172, 206  
 Simon, H.A. 12, 20, 23  
 Skorupski, J. 8, 9, 207, 217  
 Smart, J.J.C. 151  
 Steedman, I. 113, 114  
 Stegmüller, W. 13, 20, 24  
 Suchanek, A. 20, 22  
 Sugden, R. 5, 54, 56, 61, 71, 73, 205, 217  
 Suzumura, K. 66, 67, 71-73, 83, 84, 151
- Taylor, M. 171, 172, 189  
 Trapp, R.W. 113, 114  
 Trilling, L. 170, 172
- Ullmann-Margalit, E. 39, 41, 188, 190  
 Urbach, P. 13, 20, 24
- Vanberg, V. 17-19, 21, 22, 24  
 Voss, T. 21-23, 152, 188-190  
 Vossenkuhl, W. 8, 9, 20, 24, 171, 172, 217
- Weale, A. 8, 217

- Weber, M. 15, 16, 20, 21, 23, 24, 125, 130, 153-155, 171, 172  
Williams, B. 151, 171, 173, 207  
Wittgenstein, L. 125-127, 130  
Worrall, J. 13, 20, 24  
Wright, G.H. von 171, 173
- Zahar, E.G. 13, 20, 24

# Sachregister

- Accountability 8, 175-180
- Agent-neutrality 195
- All things considered preference (ATC preference) 77-80
- Assurance game (AG) 119, 141, 142, 147
- Axiomatic theory 82
  
- Bargaining 32, 33, 39-41, 43-46, 49, 50, 55, 57, 58, 60, 61, 80, 82, 114
- Battle of the Sexes 50-52
- Bayesian Rationality 91, 92, 204
- Bayesianism 88
  
- Categorical Imperative 121, 128
- Co-ordination 30, 50, 53, 56
- Common belief 47, 57
- Common knowledge 28, 29, 31-33, 36-39, 47, 50, 52, 54
- Condition  $\alpha$  75, 76, 83
- Consensus 85, 86, 89-99, 113, 176
- Consequentialism 29, 30, 32, 151
- Constitutive rules 125, 126
- Contractarian ethics 3, 115-117, 119-123, 128
- Contractarianism 43, 44, 55, 57, 59, 60
- Convention 44, 54-56, 60, 77
- Cooperation 30, 31, 37-40, 59, 60, 85, 152, 172, 189, 190
  
- Deliberation 6, 75-77, 79, 80-83, 192, 198
- Difference Principle 194, 196, 197
- Diskontparameter 159, 160, 171, 185-187
- Dispositionen 5, 17-19, 139
- Division game 46, 49, 50, 54-56
- Dominance relation 81-84
  
- Eigennutz 134
- Eingeschränkte Maximierer 3
- Entscheidungstheorie 131, 138, 151, 172
- Equilibrium 29, 48, 49, 53, 114
- Expected utility 32, 47, 51, 52, 87-91, 96, 117, 118, 151, 197, 203, 207
- Externe Gründe 155, 156, 171

- Fairness 33, 45, 57, 58, 156  
 Free-rider problem 56, 126  
 Free-riding 117, 118, 122, 123
- Game theory 39, 49, 51, 52, 58, 76, 82, 125, 126  
 Gefangenendilemma 8, 18, 19, 113, 114, 139-144, 146, 148, 157, 158, 190
- Heuristik 5, 12-14, 18-20  
 Homo oeconomicus 2, 12
- Impartiality 43-45, 55, 57, 193, 194, 195, 198  
 Individualism 128  
 Individualismus 1, 4  
 Initial position 44, 55, 57, 58  
 Interne Gründe 155  
 Kollektive Irrationalität 148, 149
- Kategorischer Imperativ 131, 155, 156, 165, 183  
 Klugheit 1, 4, 5, 8, 114, 170, 187, 188, 190  
 Kohärenzaxiome 134, 140, 150  
 Kohärenzpostulate 7, 137  
 Konstitutionelle Entscheidungen 17, 18  
 Koordinationsprobleme 18  
 Kosten 1, 2, 5, 11, 12, 14, 16, 17, 19, 21, 167, 168
- Leximin 194, 195, 198  
 Lorenz curves 79
- Maximax 65, 117  
 Maximin 59, 65, 117, 198  
 Meta preference 75, 80-84  
 Metapräferenz-Theorie 140  
 Moral choice 19  
 Moral psychology 115, 121  
 Moralische Dispositionen 5, 18, 19  
 Moralische Konflikte 153, 162, 164-168, 171,  
 Moralische Normen 17, 18, 22  
 Moralische Wahl 153-156, 162-164, 166-169  
 Moralphilosophie 4, 17, 19, 134  
 Mutual expectations 44
- Narrow Borda Score 78

- Nash equilibrium 48, 49, 51, 53-55  
 Natural law 44  
 Natural rights 45, 117  
 Nutzenfunktion 7, 101, 102, 106, 109, 136, 137, 183, 184  
 Nutzentheorie 134  
  
 Original position 58-60, 197  
 Other Regarding-Game 141  
  
 Pareto Principle 194  
 Pareto Unanimity 7, 85, 90-98  
 Pareto-optimality 32, 33  
 Präferenzen 2, 6-8, 19, 65, 67, 68, 70, 71, 102, 104, 106, 107, 114, 134-137, 139-151, 155, 156, 184, 185, 187-189  
 Präferenzordnung 64-67, 141, 144  
 Präferenzrelation 7, 70, 101, 135-137, 143, 146, 150  
 Praktische Vernunft 4, 5, 7, 8, 131, 137-140, 150  
 Praktischer Syllogismus 131  
 Prisoner's Dilemma 40, 49, 50, 52, 86, 118-120  
 Prudence 36, 198, 199, 201, 202, 204, 206  
 Public Good 25, 32, 35  
 Pure choice approach 75, 76  
  
 Rational agent 27, 28, 30, 32-37, 75, 77, 78, 85-88, 92, 116, 119-123, 128  
 Rational choice 19, 24, 26, 28-31, 33, 40, 44, 47, 51, 54, 57, 60, 75-78, 80-84, 88, 113, 114, 116, 120, 126, 151, 197, 205  
 Rational-Choice-Modell (RCM) 11, 14, 16, 17, 19  
 Rational determinacy 48, 49, 51, 52  
 Rational interaction 29, 32, 40  
 Rational self-interest 35, 128  
 Rationalismus 12, 13, 15, 23  
 Rationalität 1, 4, 7, 11-16, 19, 20, 22-24, 101-103, 105, 106, 108, 109, 112, 151, 153-155, 157, 160, 161, 171, 172, 183  
     kollektive - 8, 138, 139, 142-151  
     moralische - 5, 8, 18, 155, 156, 162, 170  
     ökonomische - 5, 7, 131, 133, 137, 139, 140, 150, 154-157, 160-162, 166, 168-170  
     praktische - 7, 131-134, 136, 138, 140, 141, 143, 144  
     Typen von - 11, 20, 153-155, 160-162, 166, 169, 170  
 Rationalitätspostulate 135, 183  
 Rationality 22-24, 28, 29, 31, 32, 35, 37-40, 45-47, 51, 52, 55, 57, 58, 60, 77, 78, 80, 81, 85, 86, 89, 98, 113, 117, 120, 122, 129, 130, 181, 207  
     collective - 32

- expressive - 123, 126
- instrumental - 121, 124, 192
- interactive - 32, 36, 38
- normative - 175
- Reflective equilibrium 43
- Regelutilitarismus 21, 183
- Regulative rules 125
- Revealed preference analysis 76
- Rule of the Road 53-55
  
- Saliency 56, 57
- Selbstinteresse 5, 8, 15, 23, 154-157, 160, 162, 168, 169
- Separability 5, 29-32, 35, 39
- Single objective approach 75, 76
- Social Choice 63, 71, 73, 84, 86, 89, 98, 100, 176
- Social welfare function 79, 89
- Spielform 6, 71, 72
- Spieltheorie 2, 4, 131, 132
- State of nature 43, 44, 57, 122
- Strategic interaction 30, 53
  
- Trade-Offs and multiple objectives approach (TAMO) 75, 80
- Tentative choice 82, 83
- Threshold Justice 192, 195, 196, 198, 199, 203-205
- Tit-for-Tat 18, 119, 158-160, 171, 185, 186
  
- Uneingeschränkte Maximierer 3
- Uniquely rational solution 46-52, 56, 57
- Universalisability 195, 198
- Universality 43
- Unmöglichkeitsergebnisse 63
- Utilitarianism 35, 59, 114, 115, 151, 192, 193, 196, 206, 207
  - aggregate - 193-195
  - classical - 191, 192, 198
  - philosophical - 193-195, 198, 204, 206
- Utilitarismus 8, 131, 133, 172
  
- Veil of ignorance 58, 59, 122
- Verhaltenserwartungen 17
- Vernünftige Wahl 8, 153, 161, 166, 169, 170
- Verteilungsgerechtigkeit 114, 134
  
- Weak Axiom 76

- Wertrationalität 125, 153
- Wirtschaftsliberalismus 133
- Wohlfahrtsfunktion 112, 133
- Wohlfahrtsökonomie 112
  
- Zweckrationalität 1, 125