

Julian Nida-Rümelin

Politische  
Philosophie der  
Gegenwart

Rationalität und  
politische Ordnung

W. Fink

UTB



UTB 3242

## **Eine Arbeitsgemeinschaft der Verlage**

Böhlau Verlag · Köln · Weimar · Wien  
Verlag Barbara Budrich · Opladen · Farmington Hills  
facultas.wuv · Wien  
Wilhelm Fink · München  
A. Francke Verlag · Tübingen und Basel  
Haupt Verlag · Bern · Stuttgart · Wien  
Julius Klinkhardt Verlagsbuchhandlung · Bad Heilbrunn  
Lucius & Lucius Verlagsgesellschaft · Stuttgart  
Mohr Siebeck · Tübingen  
C. F. Müller Verlag · Heidelberg  
Orell Füssli Verlag · Zürich  
Verlag Recht und Wirtschaft · Frankfurt am Main  
Ernst Reinhardt Verlag · München · Basel  
Ferdinand Schöningh · Paderborn · München · Wien · Zürich  
Eugen Ulmer Verlag · Stuttgart  
UVK Verlagsgesellschaft · Konstanz  
Vandenhoeck & Ruprecht · Göttingen  
vdf Hochschulverlag AG an der ETH Zürich

# Grundzüge der Politikwissenschaft

---

Herausgegeben von Mir A. Ferdowsi

Bisher erschienen:

Christian Schwaabe: *Politische Theorie 1.*

*Von Platon bis Locke* UTB 2931

Christian Schwaabe: *Politische Theorie 2.*

*Von Rousseau bis Rawls* UTB 2932

Petra Stykow: *Vergleich politischer Systeme*, UTB 2933

Maren Becker/Stefanie John/Stefan A. Schirm:

*Globalisierung und Global Governance* UTB 2965

C. Daase/J. MacKenzie/N. Mossbauer/P. Stykow:

*Politikwissenschaftliche Arbeitstechniken* UTB 3137

Julian Nida-Rümelin

# Politische Philosophie der Gegenwart

Rationalität und politische Ordnung

Unter Mitarbeit von Christine Bratu  
und Thomas Schmidt

Wilhelm Fink

Der Herausgeber:

Mir A. Ferdowsi, Dr. phil. habil., apl. Professor für Politikwissenschaft und Akademischer Direktor am Geschwister-Scholl-Institut für Politische Wissenschaft der Universität München. Veröffentlichungen u.a.: (Hg.): Internationale Politik im 21. Jahrhundert (2002); zus. mit Volker Matthies (Hg.): Den Frieden gewinnen. Zur Konsolidierung von Friedensprozessen in Nachkriegsgesellschaften (2003); zus. mit Dietmar Herz und Marc Schattenmann (Hg.): Von himmlischer Ordnung und weltlichen Probleme (2003); (Hg.): Afrika – ein verlorener Kontinent? (2004); (Hg.): Weltprobleme, 6. Neuausgabe (2007).

Der Autor:

Julian Nida-Rümelin, 1993-2003 Lehrstuhl für Philosophie an der Universität Göttingen, 2004 bis 2009 Lehrstuhl für Politische Theorie und Philosophie am Geschwister-Scholl-Institut für Politikwissenschaft, seit Sommersemester 2009 Lehrstuhl für Philosophie an der Ludwig-Maximilians-Universität München. Mitglied der Berlin-Brandenburgischen Akademie der Wissenschaften und der Europäischen Akademie der Wissenschaften und Künste, Kuratoriumsvorsitzender des Deutschen Studienpreises. Seit Januar 2009 Präsident der Deutschen Gesellschaft für Philosophie.

*Wichtigste Buchpublikationen:* Kritik des Konsequentialismus (1993), Logik kollektiver Entscheidungen (1994, zusammen mit Lucian Kern), Economic Rationality and Practical Reason (1997), Demokratie als Kooperation (1999), Strukturelle Rationalität (2001), Ethische Essays (2002), Handbuch Angewandte Ethik (2005), Über menschliche Freiheit (2005), Demokratie und Wahrheit (2006), Philosophie und Lebensform (2009).

Bibliografische Information der Deutschen Nationalbibliothek

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

© 2009 Wilhelm Fink GmbH & Co. Verlags-KG  
(Wilhelm Fink GmbH & Co. Verlags-KG, Jühenplatz 1, D-33098 Paderborn)  
ISBN 978-3-7705-4850-7

Internet: [www.fink.de](http://www.fink.de)

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

Printed in Germany.

Einbandgestaltung: Atelier Reichert, Stuttgart  
Herstellung: Ferdinand Schöningh, Paderborn

**UTB-Bestellnummer: 978-3-8252-3242-9**

## Vorwort des Herausgebers

Man mag es begrüßen oder es bedauern, unbestreitbar ist aber, dass mit der Vollendung des Bologna-Prozesses und der flächendeckenden Einführung von Bachelor-Studiengängen sich nicht nur die Hochschullandschaft grundlegend verändern wird, sondern dass damit auch wir, die Hochschullehrer, vor gewaltigen Herausforderungen in der Lehre stehen. Nicht unerheblich wird auch die Last sein, die auf die Studierenden zukommt. Denn es bedarf eines großen Engagements und eines umfangreicheren Zeitaufwandes als bislang, um sich in der relativ kurzen Zeit von vier bis fünf Semestern ein Basis-Wissen des Faches anzueignen und die vielen obligatorischen Module auch zu bestehen bzw. die entsprechenden ECTS-Punkte zu erwerben.

Vor allem die Tendenz zur „Verschlankung“ des Studiums erfordert übersichtliche, aber nicht weniger umfassende und fachlich fundierte Lehrbücher. Die Reihe „Grundzüge der Politikwissenschaft“, deren einzelnen Bände sich thematisch an die geläufigen Module des Bachelor-Studiums orientieren, hat sich zum Ziel gesetzt, Lehrbücher neuen Typs zu konzipieren, die – von exzellenten Vertretern des Faches verfasst – in knapp ausgearbeiteter Form Einführungen in die Studieninhalte bieten. Sie sollen den Studierenden dazu verhelfen, sich veranstaltungsbegleitend und durch Selbststudium in ein für sie neues Fachgebiet einzuarbeiten. Der Text ist lesefreundlich und konzentriert sich auf die wesentlichen Informationen des jeweiligen Themenbereichs. Dadurch erhalten die Studierenden einen schnellen und umfassenden Überblick und eine Grundlage für weiterführende, vertiefende Studien.

Der Herausgeber, die Autorinnen und der Autor hoffen, mit dem vorliegenden Band zum schnellen und erfolgreichen Studienablauf beizutragen.

Mir A. Ferdowsi

# Inhaltsverzeichnis

Vorwort . . . . .	9
1 Individuelle Rationalität . . . . .	13
1.1 Das Grundmodell rationaler Entscheidung . . . . .	15
1.2 Nutzen und Wahrscheinlichkeit . . . . .	21
1.3 Das Kooperationsdilemma . . . . .	36
2 Kollektive Rationalität . . . . .	45
2.1 Das Problem kollektiver politischer Entscheidungsfindung . . . . .	47
2.2 Der Konflikt von Liberalität und Rationalität . . . . .	58
2.3 Strukturelle Rationalität . . . . .	66
3 Klassische Paradigmen politischer Ordnung . . . . .	81
3.1 Kontraktualismus allgemein . . . . .	83
3.2 Thomas Hobbes – Rationales Eigeninteresse . . . . .	96
3.3 John Locke – Wahrung der Individualrechte . . . . .	114
3.4 Jean-Jacques Rousseau – Republik als sittliche Körperschaft . . . . .	125
3.5 Immanuel Kant – Politische Ordnung als universelle moralische Pflicht . . . . .	138
4 Zeitgenössische Paradigmen politischer Ordnung . . . . .	151
4.1 David Gauthier – Eigeninteressierter Liberalismus . . . . .	153
4.2 Robert Nozick – Eigentumsrechtbasierter Libertarismus . . . . .	169
4.3 John Rawls – Universalistischer Liberalismus . . . . .	187
5 Philosophie der Demokratie . . . . .	203
5.1 Demokratie als Kooperation . . . . .	205
5.2 Demokratie und Wahrheit . . . . .	221
5.3 Konzeptionen der politischen Philosophie und Positionen politischer Praxis . . . . .	236
Personenregister . . . . .	243





## Vorwort

Als mich Herr Professor Ferdowski einlud, zu der neuen, von ihm konzipierten Reihe im Fink-Verlag, die sich speziell an die BA-Studierenden der politischen Wissenschaft richtet, einen Band zur Politischen Philosophie der Gegenwart beizusteuern, habe ich gerne zugesagt. Das BA-Studium ist in höherem Maße verschult als der auslaufende Magister-Studiengang und es bedarf daher übersichtlicher Darstellungen der verschiedenen Teilbereiche der politischen Wissenschaft. Die Lektüre dieser Bände kann den Besuch von Vorlesungen und – gerade im Teilbereich Politische Theorie und Philosophie – das Studium der Quellen-Texte natürlich nicht ersetzen. Vielmehr sollen sie eine erste Orientierung geben, die dann zum eigenständigen Weiterstudium anregt.

Der vorliegende Band beruht auf Vorlesungen, die ich an den Universitäten Göttingen, München und St. Gallen gehalten habe, vor allem aber auf meiner Münchner Vorlesung zum Thema „Rationalität und politische Ordnung“. Die Darstellung ist insgesamt *systematisch*, nicht historisch ausgerichtet: Es geht um das Verhältnis von individueller Entscheidungsrationaleität einerseits und der Institutionalisierung einer politischen Ordnung andererseits. Entsprechend ist der Band in fünf Teile gegliedert:

Ein erster befasst sich mit individueller Entscheidungsrationaleität und dem für Politische Theorie und Philosophie zentralen Problem der Kooperation (Teil 1). Ein zweiter stellt die wichtigsten Ergebnisse der sog. *collective choice*-Theorie (die im Englischen alternativ auch als *social-* oder *public choice-* *theory* bezeichnet wird) dar – ein Zweig der rationalen Entscheidungstheorie, der auf dem europäischen Kontinent in der Regel nur schwach rezipiert wird (Teil 2). Ich bin jedoch der Auffassung, dass Politische Philosophie ohne Kenntnis der wichtigsten Resultate der Logik kollektiver Entscheidungen nicht studiert werden kann.

Diese Überlegungen zur Rationalitätstheorie bilden den Grundstein für das weitere Nachdenken über politische Ordnung. Denn die moderne Politische Philosophie beginnt mit der Idee, politische Ordnung auf die individuelle Rationalität ihrer Bürger zu gründen. Auf diese Art und Weise wird das kontraktualistische Programm der normativen Politischen Philosophie initiiert. Der dritte Teil des vorliegenden Bandes führt in dieses Paradigma ein, indem er den Kontrak-

tualismus zuerst allgemein vorstellt (3.1) und dann anhand der vier klassischen Paradigmen des vertragstheoretischen Denkens diskutiert, die auch heute noch für das Verhältnis von Rationalität und politischer Ordnung von zentraler Bedeutung sind: So gründet Thomas Hobbes die politische Ordnung auf dem rationalen Eigeninteresse der Individuen, das sie nach einem Ausgang aus dem Krieg aller gegen alle im Naturzustand suchen lässt (3.2). John Locke hält zwar am rationalen Eigeninteresse als zentraler Begründungsinstanz der politischen Ordnung fest, fügt aber die Idee unveräußerlicher Menschenrechte – das Recht auf Leben, körperliche Unversehrtheit und rechtmäßig erworbenes Eigentum – hinzu. Die rechtlich verfasste staatliche Ordnung hat diese Grundfreiheiten aller Menschen zu sichern (3.3). Jean-Jacques Rousseau versteht die Republik dagegen als sittliche Körperschaft der Bürger, die sich gemeinsam Gesetze geben, die ihre bürgerliche Freiheit und Gleichheit sichern und der Verfolgung der einzelnen Privatinteressen Grenzen setzen (3.4). Immanuel Kant schließlich entwickelt eine Konzeption des ursprünglichen Vertrages als Prüfstein aller Gesetzgebung und verbindet damit ethische und politische Philosophie (3.5).

Mit John Rawls setzt in den 1970er Jahren eine Renaissance des kontraktualistischen politischen Denkens ein, die bis heute anhält und den klassischen Paradigmen von Hobbes, Locke, Rousseau und Kant eine neue Aktualität verschafft hat. Denn die unterschiedlichen zeitgenössischen Konzeptionen des Kontraktualismus lassen sich als Fortentwicklung dieser Paradigmen verstehen und unter Rückgriff auf die klassischen Autoren strukturieren und interpretieren. Dies wird im vierten Teil dieses Bandes deutlich: David Gauthiers Konzeption ist in ihrem Ausgang hobbesianisch (ursprünglich ging sie aus einer Rekonstruktion der Hobbes'schen Theorie hervor), nimmt aber auch ein wichtiges Element des Locke'schen Paradigmas auf. Zudem ist sie besonders eng mit der Entscheidungs- und Spieltheorie verbunden und kann als die rationalitätstheoretisch am weitesten durchdachte Konzeption der Politischen Philosophie der Gegenwart gelten (4.1). Robert Nozick argumentiert dagegen ausschließlich in Locke'schen Begriffen und gründet politische Ordnung auf die ursprünglichen Eigentumsrechte der Individuen, weshalb diese Konzeption und ihre Nachfolger als libertäre Theorie bezeichnet werden. Der Liberalismus hält staatliche Umverteilung, etwa über Steuern, für illegitim und stellt die freiwillige Übereinkunft ins Zentrum politischer Ordnung (4.2). Die elaborierteste und komplexeste Form des zeitgenössischen Kontraktualismus ist jedoch die seines Initiators

John Rawls' (4.3). Sie integriert mehrere der klassische Paradigmen (wenn auch mit einem Schwerpunkt auf Kant und Rousseau), enthält konsequentialistische ebenso wie deontologische Elemente und beansprucht eine Gerechtigkeitskonzeption entwickelt zu haben, die mit einem breiten Spektrum unterschiedlicher moralischer Auffassungen und Lebensformen vereinbar ist und dem tatsächlichen öffentlichen Vernunftgebrauch westlicher Demokratien entspricht.

Im fünften und letzten Teil diskutiere ich einige Implikationen, die aus den vorherigen Teilen für die Philosophie der Demokratie gezogen werden können: Welche Rolle spielt Kooperation für Demokratie und wird diese durch Wahrheitsansprüche gefährdet (5.1 und 5.2)? Da die Paradigmen der zeitgenössischen Politischen Philosophie in einem auffälligen Gegensatz zu den programmatischen Positionen der politischen Praxis stehen, schließt der Band mit einigen Erläuterungen zu diesem Verhältnis (5.3). Man kann durchaus sagen, dass sich die politische Philosophie von den ideologischen Voreingenommenheiten der Nachkriegszeit gelöst hat und in einer eigenen Souveränität die politischen Frontstellungen überwinden konnte. Manches deutet darauf hin, dass es unterdessen auch einen Einfluss von Seiten dieser entideologisierten Philosophie auf die Programmatik politischer Parteien in den westlichen Demokratien gibt.

Die rationalitätstheoretischen Teile dieses Bandes stützten sich auf meine Monographie *Entscheidungstheorie und Ethik* (München, 2005), auf die zusammen mit Lucian Kern verfasste *Logik kollektiver Entscheidungen* (München/Wien, 1994) und vor allem auf das zusammen mit Thomas Schmidt (der mittlerweile an der Humboldt-Universität zu Berlin lehrt, zum Zeitpunkt der Niederschrift aber mein wissenschaftlicher Assistent an der Universität Göttingen war) verfasste Buch *Rationalität in der praktischen Philosophie* (Berlin, 2000). Letzteres ist unterdessen vergriffen und nur noch über die Wissenschaftliche Buchgesellschaft erhältlich. Die Teile 3 und 4 wurden von meiner wissenschaftlichen Mitarbeiterin, Frau Christine Bratu, auf der Grundlage ihrer Vorlesungsmitschrift erstellt. Das Kapitel 5.1 ist die leicht gekürzte Fassung eines gleichnamigen Aufsatzes von mir, der dem Sammelband *Demokratie als Kooperation* (Frankfurt a.M., 1999) seinen Titel verliehen hat. Das Kapitel 5.2 befasst sich mit der Rolle normativer und empirischer Wahrheitsansprüche in der modernen Demokratie. Das Verhältnis von Demokratie und Wahrheit hatte ich bewusst als Thema meiner Antrittsvorlesung 2004 (nach meinem Wechsel von der Universität Göttingen nach München) gewählt, weil ich es für die zentralste aktuelle Problematik

der Politischen Philosophie halte. In dem 2006 erschienenen Buch *Demokratie und Wahrheit* sind weitere Aspekte dieser Problematik behandelt.

Der Band befasst sich mit einer zentralen Frage der modernen Politischen Philosophie, nämlich dem Verhältnis von individueller Rationalität und politischer Ordnung. Er erhebt keineswegs den Anspruch einen Überblick über das gesamte Spektrum der zeitgenössischen politischen Philosophie zu geben. Diese Selbstbeschränkung hat den Vorteil, dass die Argumente systematisch entwickelt werden können, aber den Nachteil, dass vieles ausgeblendet bleiben muss. Der von Christian Schwaabe in der gleichen Reihe verfasste Band, *Politische Theorie 2. Von Rousseau bis Rawls* behandelt insbesondere das Konzept der deliberativen Demokratie (für welches paradigmatisch Jürgen Habermas steht), die Kritische Theorie sowie die kommunitaristische Kritik am Rawls'schen Liberalismus und vervollständigt damit das Spektrum der politischen Theorie und Philosophie der Gegenwart.

# 1 Individuelle Rationalität



# 1.1 Das Grundmodell rationaler Entscheidung

## Eine erste Definition individueller Rationalität

Täglich führen wir eine Vielzahl von Handlungen aus: Wir stehen auf, frühstücken und wenden uns sodann verschiedenen Tätigkeiten zu. Dabei gehen wir fast immer davon aus, dass wir eigentlich gerade etwas anderes hätten tun können als das, was wir tatsächlich gemacht haben. Oft sind wir sogar der Auffassung, dass eine der möglichen Alternativen, die uns offen standen, in einem zu klärenden Sinne *besser* ist als die Handlung, die wir gewählt haben: Es wäre ratsam gewesen, sich anders zu *entscheiden*. Doch offensichtlich unterscheiden sich Entscheidungssituationen in vielerlei Hinsicht. Entsprechend gibt es sehr heterogene Vorstellungen darüber, welche Maßstäbe wir anlegen, wenn wir beurteilen, ob eine Entscheidung gut oder schlecht war. In der *Entscheidungstheorie*, deren „klassisches“ Grundmodell im Folgenden dargestellt werden soll, versucht man, allgemeine Begriffe und Kriterien zu entwickeln, um zu beurteilen, ob eine Entscheidung individuell vernünftig ist oder nicht.

Ein relativ direktes Verfahren, sich an die für die Bewertung von Entscheidungen relevanten Aspekte „heranzutasten“, ist zu fragen, wie man Entscheidungen von anderen Personen *kritisiert*. So wird häufig angeführt, dass die gewählte Handlung die subjektiven, d.h. von der Person angestrebten Ziele nicht in optimaler Weise befördert. Ein solcher Vorwurf wird in Sätzen wie „Wenn Du dies und jenes erreichen wolltest, hättest Du besser das und das getan“ geäußert. Selbstverständlich ist dies nur eine von vielen verschiedenen Arten, Entscheidungen zu kritisieren. Das *rational choice*-Paradigma (als dessen „formaler Ableger“ die klassische Entscheidungstheorie häufig angesehen wird) geht davon aus, dass dieser Typ von Entscheidungskritik bzw. -rechtfertigung zentral ist. Präziser gefasst lautet das Credo: Eine Handlung einer Person *P* ist genau dann rational, wenn sie aus der Sicht von *P* (d.h. mit Bezug auf *P*'s Überzeugungen) die Ziele von *P* in optimaler Weise vorantreibt bzw. verwirklicht.

Dieser Kern des *rational choice*-Paradigmas wird im Grundmodell der formalen Entscheidungstheorie präzisiert (v. Neumann/ Morgenstern, 1947; Nida-Rümelin, 2005). In Teil 1 dieses Bandes werden die Grundbegriffe dieses Modells zusammengetragen und erläutert. Dabei wird in Kapitel 1.1 eine intuitive Näherung dargestellt, die in den Kapiteln 1.2 und 1.3 formal präzisiert werden soll.

## Das Dominanz-Prinzip als Entscheidungskriterium

Beginnen wir mit einem Beispiel: Eine militärische Großmacht hat ein beträchtliches nukleares Waffenarsenal angesammelt und die Regierenden sehen sich den Forderungen einer aufgebrachten Öffentlichkeit nach Abrüstung ausgesetzt. Die Regierung steht nun vor zwei Optionen (Handlungen)  $h_1$  und  $h_2$  (vgl. Jeffrey, 1965: 2):

$h_1$ : Aufrüsten oder Status Quo beibehalten.

$h_2$ : Abrüsten.

Gehen wir weiterhin davon aus, dass die relevanten Unterschiede für die Konsequenzen der Entscheidung davon abhängen, ob es Krieg geben wird oder nicht. Zur Zeit des Kalten Krieges haben weite Kreise der US-amerikanischen Bevölkerung die geopolitische Lage so eingeschätzt, dass ein Krieg bei vorangegangener Aufrüstung die Auslöschung der Menschheit bedeuten würde. Ein Krieg ohne vorangegangene Aufrüstung durch die USA hätte einer verbreiteten Einschätzung nach in der Weltherrschaft des Kommunismus resultiert. Falls kein Krieg eintritt, würde die Aufrüstung zum Fortdauern des Status quo führen, während die Verbindung von Abrüsten mit einer friedlichen Koexistenz von den meisten als die wünschenswerteste Entwicklung angesehen wurde. Setzt man diese Sicht der Dinge voraus, so kann die Entscheidungssituation der Regierung durch folgende sog. *Konsequenzenmatrix* wiedergegeben werden:

	Krieg	Frieden
$h_1$	Vernichtung der Menschheit	Aufrechterhaltung des Status quo
$h_2$	Weltherrschaft des Kommunismus	Goldenes Zeitalter

Tabelle 1.1.1: Ein Abrüstungsbeispiel

Was sollte die Regierung tun? Die Entscheidung für die beste Handlung hängt davon ab, wie die möglichen Konsequenzen bewertet werden, d.h. welches die *Präferenzen* der Individuen über den Konsequenzen sind. Doch was ist mit „Präferenz“ gemeint bzw. wie haben wir einen Satz der Form „Person P zieht x gegenüber y (schwach) vor“ zu verstehen? Um die Wahrheitsbedingungen einer solchen Aussage festzulegen, gibt es zwei grundsätzlich verschie-



dene Möglichkeiten: Entweder könnte man die Person einfach fragen, ob sie  $x$  gegenüber  $y$  vorzieht. Wenn man diese Möglichkeit wählt, muss man sich allerdings vollständig auf die Aussagen der Person selbst verlassen. Da man die eigenen Annahmen über die Präferenzen der Person ausschließlich von deren Äußerungen abhängig macht, hat man keine Möglichkeit, deren Aussagen im Rekurs auf irgendetwas anderes zu kritisieren. Insbesondere würde eine derartige Auffassung Aussagen wie „Du hast zwar behauptet, dass du  $x$  gegenüber  $y$  vorziehst, aber dies widerspricht der Art und Weise, wie du nun handelst“ sinnlos machen. Dieses Problem wird im Rahmen des sog. *revealed preferences*-Ansatz („aufgedeckte Präferenzen“) vermieden, der in der orthodoxen Entscheidungstheorie sehr prominent geworden ist und vielfach kaum noch vom klassischen *rational choice*-Paradigma getrennt wird. Der Grundgedanke dieser alternativen Auffassungsart ist, individuelle Präferenzen auf beobachtbares Verhalten zurückzuführen: Bei Entscheidungen werden Präferenzen gleichsam offenbart.

Wie liegen nun die Präferenzen der einzelnen Bürger im dargestellten Beispiel? Damals wie heute bestand ein weitgehender Konsens, wonach die Weltherrschaft des Kommunismus der Vernichtung der Menschheit und dass Abrüstung bei friedlicher Koexistenz dem Status quo vorzuziehen sei (wobei sich diese Präferenzen z.B. in Demonstrationen pro Abrüstung zeigen könnten). Wenn diese Präferenzen vorausgesetzt werden, so könnte man folgendes Argument zugunsten der Alternative  $h_2$  („Abrüsten“) vorbringen:

- (1) Es gibt entweder Krieg oder Frieden.
- (2) Wenn es Krieg gibt, wäre  $h_2$  besser gewesen als  $h_1$ .
- (3) Wenn es Frieden gibt, wäre  $h_2$  besser gewesen als  $h_1$ .
- (4) Also sollte abgerüstet werden.

Hinter diesem Argument verbirgt sich eine Entscheidungsregel, die *Dominanzprinzip* genannt wird. Das Dominanzprinzip besagt, dass vernünftigerweise diejenige Handlung zu wählen ist, deren Konsequenzen *unter allen Umständen* besser sind als die aller anderen möglichen Alternativen. Doch das Problem des Dominanzprinzips ist, dass es nur dann sinnvoll anzuwenden ist, wenn die Wahrscheinlichkeiten der relevanten Umstände, die eintreten können (in diesem Fall also der Entscheidung der anderen Konfliktpartei), von der eigenen Entscheidung unabhängig sind. Doch das Vorliegen *handlungs-unabhängiger Wahrscheinlichkeiten* ist weder in obigem Beispiel noch grundsätzlich immer der Fall.

## Das Bayes'sche Entscheidungskriterium

Gesucht wird daher ein Entscheidungskriterium, das auch dann anwendbar ist, wenn die Wahrscheinlichkeiten der relevanten (die Konsequenzen der Entscheidungen beeinflussenden) Umstände von den Entscheidungen des Akteurs abhängig sind. Dabei verfügen die Akteure im Allgemeinen über Informationen über die *Wahrscheinlichkeiten* der Umstände, welche für das Eintreten bzw. Ausbleiben gewisser Konsequenzen ihrer Handlungen verantwortlich sind. Es ist jedoch wesentliches Merkmal vieler Entscheidungssituationen, dass nicht mit Sicherheit vorhergesagt werden kann, welche Konsequenzen eine Handlung haben wird (dass die Entscheidung also *unter Unsicherheit bzw. Risiko* stattfindet).

Das *Bayes'sche Entscheidungskriterium* bestimmt die Rationalität einer Entscheidung so, dass sowohl diese subjektiven Wahrscheinlichkeiten als auch die Stärken subjektiver Präferenzen relevant sind. Genauer gesagt geht das Grundmodell der klassischen Bayesianischen Entscheidungstheorie davon aus, dass eine *subjektive Wahrscheinlichkeitsfunktion* über der Menge der Umstände gegeben ist, die präzise Informationen darüber liefert, für wie wahrscheinlich der Handelnde das Eintreten jedes relevanten Umstandes hält. Darüber hinaus nimmt es an, dass eine *subjektive Nutzenfunktion* gegeben ist, die jeder möglichen Konsequenz eine (reelle) Zahl zuordnet, welche die subjektive Bewertung dieser Konsequenzen ausdrückt. Diese beiden Informationen erlauben es, für jede Handlung deren *Nutzenerwartungswert (NEW)* zu berechnen. Dieser berechnet sich als *die Summe der mit den jeweiligen Wahrscheinlichkeiten gewichteten Nutzenwerte der möglichen Konsequenzen einer Handlung*. Das *Bayes'sche Entscheidungskriterium* verlangt dann, die Handlung mit dem größten NEW zu wählen. Mit dieser Entscheidungsregel ist zudem ein Verfahren gefunden, alle Handlungen in eine Präferenzordnung zu bringen: Die Handlungsoptionen werden in der Reihenfolge ihrer NEWs angeordnet.

Um das Bayes'sche Kriterium anwenden zu können, braucht man also – neben Informationen über die Wahrscheinlichkeit des Eintretens der relevanten Umstände – auch Wissen um die Präferenzen der Handelnden, d.h. um die Nutzenwerte, die sie den möglichen Handlungskonsequenzen zuweisen. Das folgende Beispiel macht deutlich, dass es im Allgemeinen nicht sinnvoll ist, Nutzenwerte mit direkt beobachtbaren Eigenschaften der Konsequenzen zu identifizieren. So könnte man meinen, dass man die zahlenmäßigen Verhältnisse zwi-

schen *Anzahlen* von Dingen, die man unter gewissen Umständen bekommt oder abgeben muss, als Maß für das Verhältnis der entsprechenden Nutzenwerte verwenden kann: Zwei Schecks über den gleichen Geldbetrag sind doppelt so viele wie einer, und daher haben sie (vermeintlich) auch den doppelten Nutzen. Dass diese Art des Raisonierens inadäquat ist, zeigt das Beispiel eines Autobesitzers, für den es unplausibel wäre anzunehmen, dass  $n$  Autos  $n$ -mal soviel Wert sind (d.h.  $n$ -mal soviel Nutzen aufweisen) wie ein Auto. Denn je mehr Autos man besitzt, desto geringer ist der Nutzen eines weiteren. Die dieser Beobachtung zugrunde liegende Annahme ist in der Literatur als *Gesetz vom abnehmenden Grenznutzen* bekannt geworden: *Der subjektive Nutzen einer Einheit eines Gutes nimmt ab, je mehr Einheiten von demselben Gut der Person bereits zur Verfügung stehen.* Wie man den Nutzenwert einer Konsequenz stattdessen sinnvoll begrifflich fassen kann, wird im nächsten Kapitel ausführlich erörtert werden.

Bisher lässt sich Folgendes festhalten: Die Bayesianische Entscheidungstheorie repräsentiert Handlungen als Funktionen, die unter gewissen Umständen bzw. Bedingungen (über deren Eintreten oder Ausbleiben man im Allgemeinen nichts Sicheres sagen kann) bestimmte Konsequenzen haben. Weiterhin geht man davon aus, dass dem Entscheidungsträger eine subjektive Wahrscheinlichkeitsfunktion über der Menge der Umstände sowie eine subjektive Nutzenfunktion über der Menge der Konsequenzen zugeordnet werden kann. Die nach dem Bayes'schen Kriterium rationale Handlung ist diejenige, welche den Erwartungswert des subjektiven Nutzens maximiert.

## **Zur Interpretation des entscheidungstheoretischen Grundmodells**

An dieser Stelle scheint es sinnvoll, einen Bogen von dem eingangs informell charakterisierten Grundgedanken des *rational choice*-Paradigmas zu der bis hierher eingeführten begrifflichen Struktur des Grundmodells zu schlagen. Im Kern des *rational choice*-Ansatzes steht die Annahme, dass die Rationalität individueller Entscheidungen mit Bezug auf zwei Merkmale des Entscheidungsträgers zu beurteilen ist: seine subjektiven Wünsche (bzw. Ziele) und seine Meinungen darüber, was die adäquaten Mittel sind, um diese Wünsche zu befriedigen (bzw. Ziele zu erreichen). Diese beiden sog. *propositionalen Einstellungen* – die des Wünschens und die des Für-wahr-Haltens –

werden in das Bayesianische Grundmodell durch die Wünschbarkeits- bzw. die Wahrscheinlichkeitsfunktion aufgenommen. Insofern stellt die Bayesianische Entscheidungstheorie eine Formalisierung des normalsprachlichen Grundprinzips von *rational choice* dar.

Schon an dieser Stelle ist jedoch anzumerken, dass das entscheidungstheoretische Grundmodell keineswegs auf die enge, durch das *rational choice*-Paradigma nahe gelegte Interpretation festgelegt ist: Zum einen ist der Schluss, dass der Bayesianismus notwendig mit einer in jeder Hinsicht subjektivistischen Theorie praktischer Vernunft einhergeht, verfehlt. Zwar war bislang stets von *subjektiven* Nutzen- und Wahrscheinlichkeitsfunktionen die Rede, jedoch können die Kriterien dafür, was als vernünftige individuelle Einstellung gegenüber Zielen und Erwartungen des Eintretens von Umständen gelten kann, ergänzt werden durch Forderungen, die über bloß formale Konsistenzbedingungen hinausgehen. Die zweite Bemerkung schließt unmittelbar an diesen Punkt an: Oftmals wird behauptet, dass das entscheidungstheoretische Grundmodell ausschließlich auf eigeninteressiertes Handeln sinnvoll angewandt werden kann, bei dem es den Handelnden primär darum geht, ihre eigenen Ziele zu verfolgen. In der Tat war dies der ursprünglich intendierte Anwendungsbereich der Theorie. Jedoch gilt, *dass die Nutzenfunktion zunächst in keiner Weise inhaltlich interpretiert ist*. Sie kann also auch moralische Einstellungen gegenüber bestimmten Handlungskonsequenzen zum Ausdruck bringen.

---

## Literatur

---

- Joyce, James, *The Foundations of Causal Decision Theory*, Cambridge 1999.
- Kern, Lucian/ Nida-Rümelin, Julian, *Logik kollektiver Entscheidungen*, München/ Wien 1994.
- von Neumann, John/ Morgenstern, Oskar, *Theory of Games and Economic Behavior*, Princeton 1947.
- Nida-Rümelin, Julian, *Entscheidungstheorie und Ethik*, München 2005.
- Nida-Rümelin, Julian/ Schmidt, Thomas, *Rationalität in der praktischen Philosophie. Eine Einführung*, Berlin 2000 (insbesondere Kapitel 1).

## 1.2 Nutzen und Wahrscheinlichkeit

### Grundbegriffe

In diesem Kapitel soll nun der bei der Einführung des entscheidungstheoretischen Grundmodells verwendete Nutzenbegriff präzise definiert werden. Darüber hinaus wird kurz auf den Begriff der subjektiven Wahrscheinlichkeit eingegangen und schließlich das Entscheiden unter Unwissen diskutiert.

Die Theorie der Nutzenfunktion liegt seit den Arbeiten von John von Neumann und Oskar Morgenstern in axiomatischer Form vor (vgl. v. Neumann/Morgenstern, 1947). Um die Grundzüge der modernen Nutzentheorie klar herausarbeiten zu können, bietet es sich jedoch an, zunächst die Fundamente der Konzeption vorzustellen, wie sie ursprünglich von Frank Ramsey entwickelt wurden (vgl. Ramsey, 1926). Hierfür bedarf es einiger Begriffe und Bezeichnungen: Mit  $Z$  bezeichnen wir eine *Menge von Zuständen*  $z_1, z_2, \dots$ , wobei ein Zustand als eine möglichst vollständige Beschreibung der Welt aufgefasst werden kann. Aus technischen Gründen nehmen wir an, dass es nur endlich viele Zustände gibt, und entsprechend können wir eine Zustandsmenge der Form  $Z = \{z_1, z_2, \dots, z_n\}$  bilden. Je zwei verschiedene Zustände schließen sich gegenseitig aus, und die Zustandsmenge sei in dem Sinne vollständig, dass es nicht sein kann, dass der tatsächliche Zustand nicht Element von  $Z$  ist (hierdurch ist sichergestellt, dass *genau ein* Zustand in  $Z$  der ist, der wirklich der Fall ist). Teilmengen der Zustandsmenge sind *Ereignisse*, die wir mit  $A, B, \dots$  bezeichnen.  $E$  ist die Menge aller Ereignisse. Weiterhin sei eine – ebenfalls aus technischen Gründen endliche – Menge von *Konsequenzen*  $K = \{k_1, \dots, k_m\}$  gegeben. *Handlungen* sind in diesem begrifflichen Rahmen als Funktionen von  $Z$  nach  $K$  aufzufassen: Jede Handlung ordnet jedem Zustand eine Konsequenz zu.<sup>1</sup> Handlungen werden mit den Buchstaben  $f, g, h, \dots$  bezeichnet und in der *Menge der Handlungen*  $F$  zusammengefasst. Wir fordern, dass  $F$  alle logisch möglichen Funktionen von  $Z$  nach  $K$  umfasst.

Da wir zunächst nur an der Konstruktion einer Nutzenfunktion interessiert sind, gehen wir davon aus, dass die subjektive Wahr-

---

<sup>1</sup> Wenn eine Handlung  $f$  jedem Zustand eines Ereignisses  $A$  dieselbe Konsequenz  $k_1$  zuweist, so schreiben wir (wenn auch formal etwas unpräzise)  $f(A) = k_1$ .

scheinlichkeitsfunktion des Entscheidungsträgers bekannt ist (wobei wir in diesem Kapitel stets voraussetzen, dass die Wahrscheinlichkeiten der relevanten Umstände im in Kapitel 1.1 dargestellten Sinne handlungsunabhängig ist). Es sei also eine Funktion der subjektiven Wahrscheinlichkeiten  $P$  gegeben, die jedem Ereignis aus  $E$  eine reelle Zahl zuweist und die folgenden Eigenschaften hat:

- (1) Für alle  $A \in E$ :  $P(A) \geq 0$ .
- (2) Wenn  $A \cap B = \{\}$ , dann ist  $P(A \cup B) = P(A) + P(B)$ .
- (3)  $P(Z) = 1$ .<sup>2</sup>

(Aus diesen Forderungen folgt, dass  $P$  nur Werte im abgeschlossenen Intervall zwischen 0 und 1 annimmt.)

Ein einfaches Beispiel soll diese Konzepte verdeutlichen: Wir betrachten ein Würfelspiel, bei dem ein fairer Würfel einmal geworfen wird. Ein Spieler kann einen Tipp abgeben und sich hierbei zwischen den Möglichkeiten „gar nicht spielen“ ( $f$ ), „gerade Zahl“ ( $g$ ) und „Zahl kleiner drei“ ( $h$ ) entscheiden. Wenn er sich für die erste Möglichkeit entscheidet und Recht behält, bekommt er 3 Euro ausbezahlt. Im Falle der richtigen Voraussage einer Zahl kleiner als drei ist die Auszahlung 5 Euro. Wenn er gar nicht spielt, bekommt er nichts – unabhängig davon, was der Würfel zeigt. Die Menge der Zustände ist demnach  $Z = \{1, 2, 3, 4, 5, 6\}$ .  $A = \{2, 4, 6\}$  ist das Ereignis „gerade Zahl“,  $B = \{1, 2\}$  das Ereignis „Zahl kleiner 3“, und die Konsequenzmenge ist  $K = \{0 \text{ Euro}, 3 \text{ Euro}, 5 \text{ Euro}\}$ . Die Handlungen  $f$ ,  $g$  und  $h$  sind wie folgt definiert (wobei für jedes Ereignis  $A$  das Ereignis  $\neg A$  dadurch bestimmt ist, dass es genau dann eintritt, wenn  $A$  nicht eintritt):

$$f: f(Z) = f(1) = f(2) = f(3) = f(4) = f(5) = f(6) = 0 \text{ Euro.}$$

$$g: g(A) = g(2) = g(4) = g(6) = 3 \text{ Euro.}$$

$$g(\neg A) = g(1) = g(3) = g(5) = 0 \text{ Euro.}$$

$$h: h(B) = h(1) = h(2) = 5 \text{ Euro.}$$

$$h(\neg B) = h(3) = h(4) = h(5) = h(6) = 0 \text{ Euro.}$$

Nimmt man weiterhin an, dass die subjektive Wahrscheinlichkeitsfunktion der eines fairen Würfels entspricht, so ist  $P$  gegeben durch  $P(z) = 1/6$  für alle  $z \in Z$ . Insbesondere ist  $P(A) = 1/2$  und  $P(B) = 1/3$ .

<sup>2</sup> „ $\in$ “ steht für „ist Element von“; mit „ $A \cap B$ “ bezeichnet man die Schnittmenge der Mengen  $A$  und  $B$ , mit „ $A \cup B$ “ deren Vereinigungsmenge und mit „ $\{\}$ “ die leere Menge (vgl. etwa Kutschera/Breitkopf, 1985).

## Die Zuweisung von Nutzenwerten nach Ramsey

Wenden wir uns nun dem Problem der Nutzenfunktionen zu. Eine Nutzenfunktion ist eine Funktion  $u$ , die jeder Konsequenz aus  $K$  eine reelle Zahl zuordnet. Doch wie kann man einer Person eine Nutzenfunktion zuschreiben? Der erste Schritt ergibt sich aus dem Ziel, das wir über den „Umweg“ der Nutzenfunktion erreichen wollen: Wir möchten die Nutzenfunktion verwenden um mittels des Bayes'schen Kriteriums eine Präferenzordnung über den Handlungen zu gewinnen. Hieraus folgt, dass mögliche Unterschiede zwischen Nutzenfunktionen als irrelevant angesehen werden können, solange sie – bei einer beliebigen gegebenen Wahrscheinlichkeitsfunktion – dieselbe Präferenzordnung über der Menge der Handlungen ergeben. Solche Nutzenfunktionen nennt man *äquivalent*. Mit relativ geringem technischen Aufwand lässt sich der Beweis der folgenden Aussage führen: Zwei Nutzenfunktionen sind genau dann äquivalent, wenn sich die eine in die andere durch Addition einer beliebigen (u.U. auch negativen) Zahl und/oder durch Multiplikation mit einer beliebigen positiven Zahl überführen lässt. Technischer ausgedrückt:  $u_1$  und  $u_2$  sind genau dann äquivalent, wenn es reelle Zahlen  $a$  und  $b$  (mit  $a > 0$ ) gibt, so dass  $u_1 = axu_2 + b$ .

Ebenso einfach zu beweisen ist, dass auch die Addition einer beliebigen Konstanten zu allen Nutzenwerten in einer Nutzenmatrix nichts an der relativen Größe der Nutzenerwartungswerte – und damit an der Präferenzordnung – ändert. Hieraus ergibt sich die Tatsache, dass jede beliebige Nutzenfunktion äquivalent zu einer ist, deren größter Wert 1 und deren niedrigster Wert 0 ist. Durch „Verschieben“ der Nutzenfunktion (also durch Addition bzw. Subtraktion einer Konstanten) legt man die Nutzenwerte so, dass der niedrigste Wert 0 ist. Dann wird die Funktion so „gestaucht“ bzw. „gestreckt“ (d.h. mit einer positiven Konstanten multipliziert), dass ihr höchster Wert 1 ist. Unter Rückgriff auf den Begriff der Äquivalenz von Nutzenfunktionen ist es möglich, die Menge aller möglichen Nutzenfunktionen in Klassen äquivalenter Nutzenfunktionen zu zerlegen. Da es für die Analyse einer Entscheidungssituation unerheblich ist, welche der Nutzenfunktionen aus einer Äquivalenzklasse „die richtige“ ist, kann nun stets davon ausgegangen werden, dass der Nutzen der am höchsten bewerteten Konsequenz 1 und der am geringsten bewerteten Konsequenz 0 ist. (Man spricht auch von einer „Normierung der Nutzenfunktion auf 0 und 1“.)

Frank Ramsey hat einen Trick erdacht, um die Frage zu beantworten, wie der Nutzenwert einer präferentiell zwischen der besten und

der schlechtesten Konsequenz stehenden Handlung zu bestimmen ist (vgl. Ramsey, 1926): Sei  $k_1$  die beste und  $k_m$  die schlechteste Konsequenz, und sei  $k_i$  irgendeine weitere Konsequenz, die schlechter als  $k_1$ , aber besser als  $k_m$  bewertet wird. Unter der Voraussetzung, dass  $u(k_1) = 1$  und  $u(k_m) = 0$ , ist der Wert  $u(k_i)$  zu bestimmen. Das Ramsey'sche Verfahren besteht darin, der betreffenden Person sog. *Lotterien* anzubieten (dieser Begriff wird weiter unten formal präzisiert). Die Person zieht nach Voraussetzung  $k_1$  gegenüber  $k_i$  und  $k_i$  gegenüber  $k_m$  vor. Daher sollte sie auch ein Zufallsexperiment, das ihr mit einer sehr hohen Wahrscheinlichkeit  $p$  die Konsequenz  $k_1$  verspricht (und mit der geringen Wahrscheinlichkeit von  $1-p$  die Konsequenz  $k_m$ ), gegenüber der sicheren Konsequenz  $k_i$  vorziehen. Aus einem analogen Grund wird sie sich eher für  $k_i$  entscheiden als sich auf ein Zufallsexperiment einzulassen, das mit einer hohen Wahrscheinlichkeit  $q$  die Konsequenz  $k_m$  liefert und mit  $1-q$  die Konsequenz  $k_1$ . Verkleinert man die Wahrscheinlichkeiten  $p$  und  $q$  sukzessive, so ist es plausibel, dass es genau eine Wahrscheinlichkeit  $p^*$  gibt, so dass die Person indifferent zwischen der Lotterie „ $k_1$  mit  $p^*$  und  $k_m$  mit  $(1-p^*)$ “ und der sicheren Konsequenz  $k_i$  ist. Mit anderen Worten: Die Person ist indifferent zwischen den beiden Handlungen  $f$  und  $g$  in der Entscheidungssituation, die durch die in Tabelle 1.2.1 dargestellten Nutzen- und Wahrscheinlichkeitsmatrizen gegeben ist. ( $A$  und  $B$  sind die beiden Ausgänge eines beliebigen Zufallsexperiments, bei dem  $P(A) = p^*$  und  $P(B) = 1-p^*$ .)

	A	B
f	1	0
g	x	x

	A	B
f	$p^*$	$1-p^*$
g	$p^*$	$1-p^*$

Tabelle 1.2.1: Nutzen- und Auszahlungsmatrizen der Ramsey-Situation

$f$  ist die Option, die mit der Wahrscheinlichkeit  $p^*$  zu der Konsequenz  $k_1$  (und damit zu  $u(k_1) = 1$ ) und mit der Wahrscheinlichkeit von  $1-p^*$  zu  $k_m$  (und damit zu  $u(k_m) = 0$ ) führt. Die Alternative  $g$  hat mit Sicherheit die Konsequenz  $k_i$ , deren Nutzen  $x$  wir noch nicht kennen. Nun war die Wahrscheinlichkeit  $p^*$  so gewählt worden, dass die Person indifferent zwischen  $f$  und  $g$  ist. Dies bedeutet, dass die Nutzenerwartungswerte beider Handlungen gleich sind. Wir können also ansetzen:

$$1 \times p^* + 0 \times (1-p^*) = x$$