Schriften des Vereins für Socialpolitik

Band 228/II

Wirtschaftsethische Perspektiven II

Unternehmen und Organisationen
Philosophische Begründungen
Individuelle und kollektive Rationalität

Von

Wolfgang Buchholz, Ulrich Gähde, Simon Grand,
Rainer Hegselmann, Wilfried Hinsch, Margit Osterloh,
Birger P. Priddat, Regine Tiemann, Josef Wieland

Herausgegeben von

Hans G. Nutzinger

Duncker & Humblot · Berlin

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Schriften des Vereins für Socialpolitik

Gesellschaft für Wirtschafts- und Sozialwissenschaften

Neue Folge Band 228/II

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

SCHRIFTEN DES VEREINS FÜR SOCIALPOLITIK
Gesellschaft für Wirtschafts- und Sozialwissenschaften
Neue Folge Band 228/II

Wirtschaftsethische Perspektiven II

Duncker & Humblot · Berlin

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Wirtschaftsethische Perspektiven II

Unternehmen und Organisationen
Philosophische Begründungen
Individuelle und kollektive Rationalität

Von

Wolfgang Buchholz, Ulrich Gähde, Simon Grand, Rainer Hegselmann, Wilfried Hinsch, Margit Osterloh, Birger P. Priddat, Regine Tiemann, Josef Wieland

Herausgegeben von

Hans G. Nutzinger

Duncker & Humblot · Berlin

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Wirtschaftsethische Perspektiven / hrsg. von
Hans G. Nutzinger. – Berlin : Duncker und Humblot
(Schriften des Vereins für Socialpolitik, Gesellschaft für
Wirtschafts- und Sozialwissenschaften ; N. F., Bd. 228)
NE: Nutzinger, Hans G. [Hrsg.]; Gesellschaft für Wirtschafts-
und Sozialwissenschaften: Schriften des Vereins . . .

2. Unternehmen und Organisationen, philosophische Begründungen,
individuelle und kollektive Rationalität / von Wolfgang
Buchholz . . . – 1994
ISBN 3-428-08170-6
NE: Buchholz, Wolfgang

Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen
Wiedergabe und der Übersetzung, für sämtliche Beiträge vorbehalten
© 1994 Duncker & Humblot GmbH, Berlin
Fremddatenübernahme und Druck:
Berliner Buchdruckerei Union GmbH, Berlin
Printed in Germany

ISSN 0505-2777
ISBN 3-428-08170-6

Gedruckt auf alterungsbeständigem (säurefreiem) Papier
gemäß der ANSI-Norm für Bibliotheken

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Vorwort

Von *Hans G. Nutzinger,* Kassel

Der vorliegende Band enthält die überarbeiteten Beiträge der Sitzung des Ausschusses „Wirtschaftswissenschaften und Ethik" im *Verein für Socialpolitik,* die vom 10. bis 12. Februar 1994 in Osnabrück stattgefunden hat. Mein Dank gilt den Autorinnen und Autoren, die ihre Beiträge in relativ kurzer Zeit für den Druck fertiggemacht haben, und Frau Heike Frank (Verlag Duncker & Humblot) sowie Frau Elke Kalden (Gh Kassel) für ihre tatkräftige Unterstützung bei der Herstellung des Druckmanuskripts.

Die im vorliegenden Band vereinigten Aufsätze liefern einen anschaulichen Eindruck davon, wie sehr in den letzten zehn Jahren ethische Fragen als Probleme wirtschaftlicher Praxis und ökonomischer Theoriebildung an Bedeutung und an Interesse gewonnen haben.[1] Darüber hinaus zeigt sich, daß der anfänglich schwierige Dialog zwischen „Ökonomen" und „Philosophen" bei der Behandlung ethischer, also moralphilosophischer, Fragestellungen zunehmend offener und fruchtbarer zu werden beginnt.

Die diesjährige Arbeitstagung des Ausschusses „Wirtschaftswissenschaften und Ethik" — und damit auch der vorliegende Tagungsband — sind durch drei deutlich erkennbare thematische Schwerpunkte geprägt, die in gewisser Weise einen repräsentativen Querschnitt aus der Vielfalt der gegenwärtigen wirtschafts- und unternehmensethischen Diskussionen bieten. Den *ersten Schwerpunkt* bilden *unternehmensethische Fragestellungen* im weiteren Sinne. Josef Wieland (Münster) erörtert, ausgehend von US-amerikanischen Erfahrungen, Formen der Insti-

---

[1] In diesem Zeitraum hat der zunächst temporäre, seit 1989 ständige Ausschuß „Wirtschaftswissenschaften und Ethik" im „Verein für Socialpolitik" seine Arbeit in bisher fünf Bänden dokumentiert:
Enderle, Georges (Hrsg.): Ethik und Wirtschaftswissenschaft. Schriften des Vereins für Socialpolitik N. F. Bd. 147, Berlin 1985;
Hesse, Helmut (Hrsg.): Wirtschaftswissenschaft und Ethik. Schriften des Vereins für Socialpolitik N. F. Bd. 171, Berlin 1988, 2. Aufl. 1989;
Homann, Karl (Hrsg.): Aktuelle Probleme der Wirtschaftsethik. Schriften des Vereins für Socialpolitik N. F. Bd. 211, Berlin 1992;
Homann, Karl (Hrsg.): Wirtschaftsethische Perspektiven des Drogenproblems. Schriften des Vereins für Socialpolitik N. F. Bd. 217, Berlin 1992;
Homann, Karl (Hrsg.): Wirtschaftsethische Perspektiven I: Theorie, Ordnungsfragen, Internationale Institutionen. Schriften des Vereins für Socialpolitik N. F. 228 / I, Berlin 1994.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

tutionalisierung von Moral in der Unternehmung, die er sodann zum Anlaß für grundsätzliche theoretische Reflexionen über die unternehmensethische Problematik nimmt. Im Zentrum seines Ansatzes steht die Her- und Sicherstellung von Kooperationschancen als *ökonomische* Erklärung der Unternehmensethik, gerade angesichts der Tatsache, daß Kooperation einerseits konstitutiv für Entstehung und Funktion des Unternehmens, andererseits aber auch konstitutiv anfällig für die verschiedensten Formen des „moral hazard" ist. Trotz dieses „ökonomischen Ansatzes" warnt Wieland überzeugend davor, Ethikprogramme als „schlicht gewinnmaximierende Managementinstrumente" zu begreifen, da ein rein instrumenteller Einsatz von Ethik (unter einem heimlichen „Rücknahme-Vorbehalt") mit einer wesentlichen Voraussetzung von Ethik, nämlich der Glaubwürdigkeit der Akteure, konfligieren muß — und damit würde sich die Wirkung einer derart verkürzten Unternehmensethik gerade in ihr Gegenteil verkehren. Sein Beitrag löste im Ausschuß eine intensive und aufschlußreiche Diskussion über verschiedene — mögliche und aktuelle — Funktionen von Unternehmensethik aus, deren generelles Ergebnis sich vielleicht dahingehend zusammenfassen läßt, daß Ethik trotz ihrer (hoffentlich) für das Unternehmen nützlichen Funktion nicht auf ein reines Instrument verkürzt werden kann.

Im anschließenden Beitrag von Margit Osterloh (Zürich; Mitautoren: Simon Grand und Regine Tiemann) steht der mögliche Beitrag der ökonomischen Theorie zum organisationalen Lernen im Zentrum der Betrachtung. Ausgehend von einer individualistischen Sozialtheorie präsentiert Frau Osterloh organisationales Lernen als Lernen von Individuen in Institutionen derart, daß das so erworbene Wissen unabhängig vom Wechsel der Organisationsmitglieder in der Organisation erhalten bleibt. Sie sieht den genuinen Beitrag der ökonomischen Theorie darin, daß diese in der Lage ist, wichtige begriffliche Präzisierungen zu liefern, daß sie jedoch nicht den Anspruch erheben kann, organisationales Lernen vollständig in ihre eigene Sprache zu übersetzen. Auch spieltheoretische Ansätze sind ihrer Analyse zufolge nicht in der Lage, anspruchsvollere, qualitative Prozesse des Lernens adäquat abzubilden. Als Konsequenz für die Managementwissenschaft ergibt sich daraus, daß sie die ökonomische Theorie als einen von mehreren möglichen Bezugsrahmen in ihre Analyse mit einbeziehen, sie aber auch zugleich mit anderen disziplinären Bezugsrahmen in ein hinreichend flexibles „framework" mittlerer Abstraktionsweite einbauen sollte. Dieser Aufsatz macht deutlich, daß die Unternehmung in sowohl praktisch wie auch theoretisch bedeutsamer Weise ethische Fragestellungen aufzunehmen und interdisziplinär zu bearbeiten hat, eben deswegen, weil es um die Interaktion und Kooperation von lernenden und handelnden Individuen geht.

Den *zweiten Schwerpunkt* bilden *philosophisch begründete* Beiträge zur Wirtschaftsethik. Ulrich Gähde (Bayreuth) entwickelt am Beispiel des klassischen Utilitarismus sehr grundsätzliche und für die weitere Forschung anregende Überlegungen über die Anwendung wissenschaftstheoretischer Methoden auf ethische

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Theorien; er zeigt dabei, wie trotz des unterschiedlichen Gegenstandsbereichs im Vergleich zu empirisch-deskriptiven Theorien weitgehend analoge Mechanismen bei der Verteidigung ethischer Theorien gegen Konflikte mit basalen moralischen Intuitionen wirken, wie sie im Modell von Lakatos für die „Realwissenschaften" festgestellt wurden. Dabei bedient sich Gähde, anders als Lakatos, einer „Netz-Darstellung". In diesem Kontext ergibt sich natürlich u. a. die kritische Frage, ob der Nutzenbegriff — der ja in durchaus unterschiedlicher und auch recht flexibler Art und Weise in der ökonomischen Theorie verstanden und verwendet wird — denn tatsächlich ein so wohldefiniertes Konzept darstellt, daß es zum „harten Kern" einer Theorie gerechnet werden kann.

Im zweiten philosophisch orientierten Beitrag „Der Schleier der Unwissenheit und die Idee eines übergreifenden Konsenses" untersucht Wilfried Hinsch (Münster) zunächst den Begriff der „wohlgeordneten Gesellschaft" bei John Rawls und stellt in diesem Zusammenhang die unterschiedlichen Konsensbegriffe bei Kant und bei Hobbes einander gegenüber; Hinsch zufolge sprechen Rawls' prinzipielle Überlegungen zur Begründung seiner Gerechtigkeitsprinzipien für eine kantische Interpretation des übergreifenden Konsenses. Anschließend analysiert er verschiedene Argumente für den Rawlschen „Schleier der Unwissenheit", und dies führt bei ihm u. a. zu der Folgerung, daß ein Konsens über politische Gerechtigkeitsgrundsätze durch rationale Argumentation herbeigeführt werden kann, etwa durch den Nachweis, daß sich Elemente der „original position" bei Rawls aus dem Gedanken der moralischen Rechtfertigung politischer Grundsätze unter den Bedingungen eines vernünftigen Pluralismus ergeben. Hieraus, das zeigten die anschließenden Erörterungen im Ausschuß, folgt dann vor allem die Frage nach den Möglichkeiten eines vernünftigen Streites bzw. dem vernünftigen Austragen von Meinungsverschiedenheiten im Hinblick auf verschiedene Konzeptionen des Guten.

Die drei letzten Aufsätze des Bandes befassen sich in unterschiedlicher Weise mit Fragen des *„rational choice"*, die einen dritten Schwerpunkt bilden: Zunächst präsentiert Birger P. Priddat (Witten-Herdecke) in seinem Beitrag „Rationalität, Moral und Emotion" die (ökonomische) Konzeption von Wahlhandlungen als rationale Entscheidungen für Individuen in dem Sinne, daß ihnen als Verantwortlichen die Konsequenzen ihres Handelns zugerechnet werden. Diese moralische Norm erzeugt wiederum Emotionen und Spannungen, insbesondere dann, wenn das Individuum oder seine Umwelt Wahlhandlungen als „mißglückt" betrachten (müssen). Hier kommt nach Priddat Moral in einer ganz anderen Funktion zur Geltung, nämlich als eine Möglichkeit, verletzte oder beschädigte Akteursidentität und Akteurskompetenz wieder herzustellen. In diesem Zusammenhang werden zentrale Fragen einer auch ökonomieübergreifenden Sozialwissenschaft, wie etwa der Begriff der Verantwortlichkeit oder der — tatsächlich oder auch nur vermeintlich verletzten — Akteurskompetenz, angesprochen, die im Hinblick auf den „ökonomischen Ansatz", insbesondere den *homo oeconomicus* — ein wichtiges

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Thema früherer (und sicher auch künftiger) Tagungen und Publikationen des Ausschusses — einer vertieften Behandlung bedürfen.

Einen aufgrund der formalen Analyse besonders anregenden Versuch, die spätestens seit Wilhelm von Humboldt formulierte Befürchtung der Sozialstaats-Skeptiker, derzufolge eine sozialstaatliche Daseinsvorsorge die Netzwerke solidarischer Selbsthilfe zerstöre, spieltheoretisch zu analysieren, unternimmt nun Rainer Hegselmann (Bremen). Im Rahmen eines — bereits früher schon einmal im Ausschuß präsentierten und diskutierten — „Solidaritäts-Spiels" leitet er relativ plausible Bedingungen dafür ab, unter welchen Umständen staatliche Hilfe solidaritätszerstörende Wirkungen zeitigt oder nicht. Die spieltheoretische Analyse erlaubt darüber hinaus, auch zumindest abstrakte Möglichkeiten einer Gestaltung sozialstaatlicher Maßnahmen derart aufzuzeigen, daß solidaritätszerstörende Effekte gering gehalten werden. Natürlich ist der Weg von einer derartigen relativ einfachen spieltheoretischen Erörterung zur vollen sozialen Wirklichkeit noch sehr weit, so daß unmittelbare Handlungsanweisungen aus dem Modell nicht ableitbar sind, wie Hegselmann selbst betont. Immerhin zeigt sich, daß eine abstrakte Modellierung bestimmter sozialer Situationen viel zur Schärfung zuvor wenig kontrollierter Intuition und Plausibilitätserwägungen führen kann.

Wolfgang Buchholz (Frankfurt/Oder) untersucht schließlich in seinem Referat „Das Weniger-Ist-Mehr-Paradoxon" eine ähnliche Fragestellung, nämlich das Problem strategischer Anreize, sich selber arm zu machen, um sich vor der Ausbeutung durch altruistische Neigungen anderer zu schützen, und im Zusammenhang damit die Möglichkeit von Pareto-*Verbesserungen* bei *reduzierter* Anfangsausstattung. Hierbei verwendet er theoretische Konzepte der Familienökonomie und der Public-Choice-Theorie. Dies führt nun wiederum zu der Frage, ob nicht derartige Paradoxien eher auf die Notwendigkeit verweisen, flexible, durch soziale Erfahrungen veränderbare Präferenzen in die theoretische Betrachtung mit einzubeziehen.

Die hier publizierten Beiträge nehmen auch Anregungen und Ergebnisse aus den Erörterungen des Ausschusses auf, und sie geben, so hoffen wir jedenfalls, einen anschaulichen Einblick in typische Schwerpunkte der gegenwärtigen vielfältigen Bemühungen im Beziehungsfeld von „Wirtschaftswissenschaften und Ethik". Wenn sie darüber hinaus Anregungen für weitere Forschungen auf diesem Gebiet geben, dann hat dieser Band zweifellos seinen Zweck vollkommen erfüllt.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Inhalt

I. Unternehmen und Organisationen

Organisatorische Formen der Institutionalisierung von Moral in der Unternehmung
Von *Josef Wieland,* Münster 11

Organisationales Lernen — Was kann die ökonomische Theorie dazu beitragen?
Von *Margit Osterloh/Simon Grand/Regine Tiemann,* Zürich 37

II. Philosophische Begründungen

Zur Anwendung wissenschaftstheoretischer Methoden auf ethische Theorien: Überlegungen am Beispiel des klassischen Utilitarismus
Von *Ulrich Gähde,* Bayreuth 77

Der Schleier der Unwissenheit und die Idee eines übergreifenden Konsenses
Von *Wilfried Hinsch,* Münster 103

III. Individuelle und kollektive Rationalität

Rationalität, Moral und Emotion
Von *Birger P. Priddat,* Witten 129

Humboldts Sozialstaats-Skepsis. Eine spieltheoretische Rekonstruktion
Von *Rainer Hegselmann,* Bremen 159

Das Weniger-Ist-Mehr-Paradoxon
Von *Wolfgang Buchholz,* Frankfurt (Oder) 177

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Organisatorische Formen der Institutionalisierung von Moral in der Unternehmung

Empirische Resultate, theoretische Reflexionen

Von *Josef Wieland,* Münster

Zunächst werde ich einige Ergebnisse eines empirischen Forschungsprojektes präsentieren, das sich mit Möglichkeiten der Institutionalisierung moralischer Kommunikation in der Unternehmung beschäftigte und das ich im vergangenen Jahr in den USA durchgeführt habe. Dieser Teil des Aufsatzes verfolgt ein rein deskriptives Interesse. Es schließt sich eine eher theoretisch gehaltene Erörterung der empirischen Resultate an, in der es darum gehen wird, eine Erklärung institutionalisierter Unternehmensethik aus der Perspektive der „New Economics of Organization" zu skizzieren. Abschließend werden einige Aspekte der Frage erörtert, inwiefern Unternehmensethik als Ethik durchführbar ist.

1.

Beginnen wir mit einer organisationstheoretisch gehaltenen Definition dessen, was man unter einer Institutionalisierung ethischer Entscheidungen in der Unternehmung verstehen könnte.[1]

> „It means getting ethics formally and explicity into daily business life. It means getting ethics into company policy formation at the board and top management levels of employment. It means grafting a new branch on the corporate decision tree — a branch that reads right / wrong". (*Purcell / Weber* 1979)

Diese Definition ist brauchbar, wenn klar ist, daß immer verschiedene Möglichkeiten der Institutionalisierung von Firmenmoral denkbar sind. Sie umreißt einen unternehmenspolitischen Prozeß, der verschiedene Ergebnisse und Formen, organisatorische und / oder motivationale (moral leadership, moral heroes) haben kann. Natürlich sind in der Nacht alle Katzen grau, aber in diesem Aufsatz möchte ich mich auf die organisatorischen Mittel zur Implementierung von Ethikprogrammen in Unternehmen konzentrieren.

Bevor ich zu einer detaillierten Beschreibung übergehe, muß einem Mißverständnis vorgebeugt werden. Institutionalisierte Ethikprogramme zielen in der

[1] Vgl. für die vollständige Publikation der hier vorgetragenen Ergebnisse des Forschungsprojektes *Wieland* (1993 a).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Regel nicht darauf, Konflikte im Unternehmen oder zwischen dem Unternehmen und seiner Umwelt zu befrieden oder zu lösen. Sie verwandeln auch nicht eine „untugendhafte" Firma in eine „tugendhafte". Gegen beides spricht allein schon, daß Moral in der Wirtschaft häufig die Form eines Dilemmas annimmt, bei dem nur „tragic choices" bleiben. Knappheit als Ursache und Resultat dieses Entscheidungstyps ist weiterhin ein Grund dafür, daß auch moralische Entscheidungen polemogene Eigenschaften haben. (Vgl. *Luhmann* 1989: 370)

Heute kann man vielleicht sagen, daß sich in den vergangenen zehn Jahren in den USA fünf mögliche Instrumente zur organisatorischen Umsetzung eines firmenspezifischen Wertemusters herausgebildet und in gewisser Weise auch bewährt haben. Die in den USA entwickelten formalen Mittel sind: i) Code of Ethics / Conduct, ii) Ethics Committee of the Board of Directors, iii) Corporate Ethics and Business Conduct Office, iv) Ethics Training- and Audit Program und v) Ethics Hot Line.

Diese organisatorischen Formen und ihre unternehmenspolitische Funktion sollen nun zunächst knapp charakterisiert werden.

Der *„Code of Ethics"* oder *„Code of Conduct"* beschreibt und kodifiziert das Wertesystem des Unternehmens, formuliert daraus abgeleitete Organisationsziele und gibt allgemeine Richtlinien zur Umsetzung. Die existierenden „Codes" sind in ihrer konkreten Ausgestaltung extrem unterschiedlich. Dies bezieht sich natürlich auf den Inhalt des Wertesystems, aber auch auf die formale Gestaltung. Hier eine knappe Liste der Werte, die häufig vorkommen:

Marktführerschaft, Rentabilität, Ehrlichkeit, Rechtschaffenheit, Integrität, Einhaltung der Gesetze, Produktsicherheit, Produktqualität, Sicherheit am Arbeitsplatz, Verpflichtung auf eine humanorientierte Personalpolitik, auf Kunden- und Lieferantenorientierung (Lieferanten als Teammitglieder), Informationsoffenheit, Ehrlichkeit bei Preis-, Rechnungs- und Vertragsgestaltung, Umweltschutz, Verpflichtung, der Gemeinde zu nutzen (social responsibility).

Auf die merkwürdige Mischung wirtschaftlicher, politischer und moralischer Werte werde ich in den folgenden Abschnitten noch eingehen. Organisationstheoretisch zeigen sie, daß Unternehmen polylinguale Systeme sind. Der Punkt, den ich hier machen möchte, ist, daß ethische Unternehmensrichtlinien nur sinnvoll und wirksam sind, wenn durch eine vorgängige gründliche Diskussion der Inhalte auf allen Ebenen der Firma geklärt wird, was die Werte und Ziele des Unternehmens sind beziehungsweise sein sollen. Die Frage lautet: Für welche Werte soll der Name der Firma stehen? Gleichfalls von großer Bedeutung ist, daß die Aussagen des Verhaltenskodex eindeutig sind und daß festgeschrieben wird, wie sie in die Praxis umzusetzen sind. Die „Stunde der Wahrheit" für alle „Codes" kommt, wenn das Unternehmen eine ernsthafte Krise durchlebt oder eine schwierige Entscheidung zu treffen hat und die Anwendung der selbstauferlegten Verhaltensverpflichtungen zu signifikanten Kosten führen könnte. Wenn der CEO (Chief

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Executive Officer) und das Board of Directors Konsequenz zeigen, ist die Durchsetzung des „Code" als ein „living document" einen Schritt vorangekommen. Wenn nicht, ist er in seiner Bedeutung radikal entwertet. Die Geschichte der amerikanischen Business-Ethics-Bewegung ist voller Beispiele hierfür, und sie gehören mittlerweile zu den „Stories", die dem involvierten Unternehmen Leben, Profil und Kontinuität verleihen. Es sind diese „Stories", die sich um die Durchsetzung des „Codes" ranken, die dem Unternehmen das ermöglichen, wonach es in einer unübersichtlichen Firmenlandschaft immer streben muß: Make a difference! 1979 hatten nach einem Fortune Survey 75 von 100 Firmen einen „Code", und eine Untersuchung aus dem Jahre 1988 zeigte, daß 76% der befragten Unternehmen über ein solches Instrument verfügten. Eine weitere Untersuchung aus dem Jahre 1991/92 ergab, daß mehr als 90% der Firmen über einen Verhaltenskodex verfügten. (Vgl. Center for Business Ethics; 1992) Zumindest auf der Ebene des leitenden Management ist der „Code" Bestandteil des Arbeitsvertrages, so daß seine Verletzung Konsequenzen haben kann.

Die Abbildung I im Anhang zeigt zur Illustration den „Code" der Firma Northrop, Los Angeles.

Das *Ethics Committee of the Board of Directors* setzt sich gewöhnlich zusammen aus je einem internen und externen Direktor des Unternehmens und dem Verantwortlichen der Abteilung für „Corporate Ethics and Business Conduct". Seine Aufgabe besteht vor allem darin, die „ethische Frage" im Unternehmen zu legitimieren durch Ansiedlung des Komitees auf höchster Ebene. Dies ist die Grundvoraussetzung jedes unternehmensethischen Konzepts: Es muß von oben nach unten aktiv betrieben werden, und „oben" meint in den USA den Chairman und den Chief Executive Officer (CEO). Dies ist die fundamentalste und unumstrittenste Lehre der amerikanischen Business-Ethics-Bewegung aus den drei Jahrzehnten Praxiserfahrung: Ohne „moral leadership" der Unternehmensspitze gibt es keine erfolgreiche Unternehmensethik. Die bewußte oder unbewußte Delegation dieser Aufgabe auf subalterne Ebenen bedeutet faktisch, die Unmöglichkeit oder die Unzulänglichkeit ihrer Erfüllung zu programmieren. Dafür gibt es im wesentlichen zwei Gründe. Erstens: Ohne eindeutige Signale, daß die Unternehmensleitung moralische Fragen ernst nimmt und für relevant hält, kann in einer Firma keine ethische Aufmerksamkeit systematisch entstehen. Zweitens: Nicht selten sind die zu klärenden Fragen so konfliktbeladen, daß einzelnen Akteuren oder subalternen Funktionen deren Lösung allein nicht zugemutet werden kann. Wenn etwa von der Zahlung beziehungsweise Nichtzahlung eines Bestechungsgeldes ein wichtiger Auftrag abhängt, dann kann über das Verhalten der Firma in diesem Fall nur auf der Ebene des „Ethics Committee of the Board of Directors" in Zusammenarbeit mit der gesamten Unternehmensspitze entschieden werden.

Ist diese Grundvoraussetzung geklärt, rücken weitere Aufgaben des Ethikkomitees in das Blickfeld:

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Da ist zuallererst die Kommunikation der ethischen Ansprüche und Entscheidungen des Unternehmens an Mitarbeiter, Kapitalgeber und externe „stakeholder", das heißt an Kunden, Lieferanten, die Gemeinde, in der das Unternehmen angesiedelt ist, und so weiter. Nur wenn die Ansprüche des Unternehmens aktiv dargelegt und betrieben werden, können die wechselseitigen Erwartungen offengelegt und zur Beachtung gebracht werden. Dies gilt für alle Ebenen und Interaktionsbeziehungen.

Hinzu kommt die Überwachung der ethisch relevanten Entscheidungen und Handlungen des Unternehmens sowie die Anfertigung von Entscheidungen in schwierigen und bedeutsamen Fällen.

Schließlich gehört zu seinen Aufgaben die Schaffung und Anwendung eines Anreiz- und Bestrafungssystems für moralisches beziehungsweise unmoralisches und der Firmenpolitik entgegenlaufendes Verhalten durch Mitglieder der Unternehmung.

An der Umsetzung der Verhaltensstandards und der Beschlüsse des Komitees arbeitet das *Corporate Ethics and Business Conduct / Practices Office*. In ihm wirkt der *Ethics Officer (EO)*. Aufgabe dieser Unternehmensabteilung ist die Entwicklung und Implementierung von Kommunikations- und Beobachtungsmechanismen, die die Umsetzung der „Codes" auf allen Ebenen des Unternehmens ermöglichen. Dazu gehören:

— die Entwicklung von Mechanismen, durch die Mitarbeiter wahrgenommene ethische Probleme ohne Furcht vor negativen Konsequenzen formulieren können,

— die Entwicklung und Durchführung von Trainingsprogrammen für alle Mitarbeiter, vor allem aber für die Gewinnung neuer „passender" Mitarbeiter,

— die Aufdeckung und Kontrolle „ethischer Grauzonen" und moralisch sensibler Arbeitsbereiche,

— die Entwicklung und Durchführung eines Auditingsystems,

— die Veröffentlichungspolitik des Unternehmens im Hinblick auf ethische Fragestellungen zu gestalten,

— und schließlich die Berichterstattung an das Ethikkomitee und den Vorstand.

Der Leiter dieser Abteilung ist in der Regel reguläres oder berichterstattendes Mitglied des Ethikkomitees, und die gesamte Abteilung ist organisatorisch dem Komitee unterstellt. Ende des Jahres 1991 ist in den USA die Ethics Officer Association (EOA) als Berufsverband der EO's gegründet worden. Mitglied können professionelle Praktiker im Bereich „Business Ethics" werden, die im Management eines Unternehmens an verantwortlicher und hochrangiger Stelle mit der Durchführung von Ethikprogrammen beschäftigt sind.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Abbildung II im Anhang zeigt die entsprechende Organisationsstruktur der Boeing Company in Seattle. Dort sind etwa 24 Mitarbeiter (bezogen auf weltweit alle unternehmensethischen Aktivitäten der Firma sind es weitaus mehr) in der Ethikabteilung beschäftigt, dessen Jahresbudget 1993 bei rund 1,3 Millionen Dollar lag.

Betrachten wir nun einige Aufgaben des „Ethics Officers" etwas genauer. Das ist zunächst das *Ethics Training* und das *Audit Program.*

Im Hinblick auf die Trainingsprogramme geht es um die Schulung ethischen Wissens, aber vor allem um die Umsetzung der Unternehmenswerte auf Abteilungsebene beziehungsweise in konkrete Fragestellungen. Die genutzten Methoden sind vielfältig: Orientierungsgespräche für neue Mitarbeiter, Workshops / Seminare mit internen und externen Fallbeispielen, Videos, Ansprachen des CEO, Erläuterung der Unternehmenspolitik, persönliches Gespräch. Im Prinzip versuchen die Firmen, diese Aufgaben mit eigenem Personal, in der Regel dem jeweiligen Vorgesetzten, zu lösen.

Die bisher entwickelten und mir bekannten Auditverfahren zielen in erster Linie auf die empirische Erhebung des „moralischen Klimas", sei es auf Abteilungsebene oder im Gesamtunternehmen. So erhebt Boeing jährlich die Werte für ein „Ethikbarometer" der jeweiligen Organisationseinheit. Fällt das Barometer, wird dies als Indiz genommen, daß unmoralisches Handeln dort wahrscheinlicher wird. Was dies im Detail bedeutet und an Konsequenzen nach sich zieht, habe ich bei anderer Gelegenheit ausführlich beschrieben und kann daher hier übergangen werden. (Vgl. *Wieland* 1993a)

Es folgt die *Ethics Hot line* oder *Open line.* Dies ist ein innerbetriebliches Telefonnetz und eine nordamerikaweite toll-free-number, über die Mitarbeiter Beschwerden oder vermutete moralische oder rechtliche Unregelmäßigkeiten anonym dem „Ethics Office" oder einem „Ethics Adviser" mitteilen können. Eine problemorientierte Auswertung der Telefonate durch die „Defense Industry Initiative on Business Ethics and Conduct (DII)" in ihrem Jahresbericht 1990 erbrachte die Schwerpunkte (auszugsweise) Arbeitsplatzfragen (49,71 %), Mißbrauch von Arbeitszeit (12,19 %), Mißbrauch von Firmen- und Kundeneigentum (9,24 %), Angebote von Geschenken, Zuwendungen, Unterhaltung (8,21 %), Probleme mit Lieferanten (7,59 %), Qualitätskontrolle (2,49 %), Vertragsgestaltung, Preispolitik (1,3 %) und Produktsicherheit (1,28 %.).

Der hohe Anteil an Arbeitsplatzfragen scheint eine Folge der relativ schwach ausgebildeten Mitarbeitervertretung in nordamerikanischen Unternehmen zu sein. Weitere Schwerpunkte sind Opportunismus der Mitarbeiter und Bestechungspraktiken. Wenn man sich klarmacht, daß die meisten Unternehmen der DII technisch hochsensible Produkte herstellen und vermarkten (z. B. Steuerungssysteme oder Triebwerke für Flugzeuge), dann kann sich selbst hinter den scheinbar geringen Nennungen für Qualitätskontrolle und Produktsicherheit die Verhinde-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

rung einer Katastrophe verbergen. Auch hierzu findet sich im Anhang als Abbildung III die Illustration der Aufgabenspezifizierung einer Ethikabteilung.

Wie relevant, das heißt, wie verbreitet sind nun die hier expemplarisch beschriebenen Mittel zur Organisation eines unternehmensethischen Prozesses in der amerikanischen Unternehmenslandschaft? Das „Center for Business Ethics" am Bentley College hat 1984/85 und 1990/91 eine empirische Untersuchung über den Stand der Institutionalisierung von Ethik in Unternehmen durchgeführt. (Vgl. *Center for Business Ethics* 1986; 1992)

Die Fragebogen wurden an die Firmen der Fortune 500-Liste der bedeutendsten Industrieunternehmen und der Fortune 500-Liste der bedeutendsten Dienstleistungsunternehmen geschickt. 1984/85 antworteten 279 Firmen, 1990/91 244 Firmen. Von diesen Firmen gaben 1984/85 79,9% und 1990/91 93% an, sie institutionalisierten Ethik in ihrem Unternehmen.

Befragt nach den Mitteln in diesem Zusammenhang gab es folgende Antworten (der erste Wert bezieht sich auf 1984/85, der zweite auf 1990/91):

Code of conduct (93,3% / 93%); Training in Ethics (44,4% / 52%); Social auditing (43,9% / 30%); Change of organisational structures (20,6% / 13%); Committees of Ethics (14% / 32%).

Die Aufgaben des Ethik-Komitees wurden wie folgt spezifiziert:

Beobachtung der ethischen Aktivitäten des Unternehmens (67,5% / 81%); Entwicklung der Unternehmenspolitik (67,5% / 70%); Beantwortung von Anfragen über Zweifelsfälle oder interne Konflikte (52,5% / 62%); Behandlung von Verletzungen des „Code" (40% / 53%).

Diese Zahlen sprechen dafür, daß sich der Institutionalisierungsprozeß moralischer Kommunikation in den Unternehmen, ungeachtet einer langanhaltenden und tiefgehenden Rezession, beschleunigt hat. Dafür spricht zunächst die Abnahme der Werte für relativ unspezifische (social auditing, Umorganisierung) und die Zunahme aller Werte für spezifische Maßnahmen. Vor allem der Zuwachs an Ethikkomitees und der Aktivitätszuwachs dieser Einrichtungen weisen in diese Richtung. Hier noch einige Zahlen, die den Charakter der Ethikkomitees und deren Tätigkeit deutlicher werden lassen: 1984/85 waren 60% der Mitglieder der Komitees „Executive Officer", 45% kamen aus dem „Board of Directors", 23% waren Manager, 8% der Firmen hatten externe Persönlichkeiten in ihr Komitee berufen und weitere 8% waren Mitarbeiter unterhalb der Managerebene.

Jeweils über 80% der Unternehmen, die einen „Code" haben, hatten ebenfalls Sanktionen spezifiziert, um dessen Einhaltung sicherzustellen. Die Skala reicht von Entlassungen, Rückstufungen, Verweisen bis hin zu Geldstrafen. Von besonderer Bedeutung ist, daß positive Anreize ausschließlich symbolischer Art sind.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

2.

In der Neuen Institutionenökonomik werden seit einigen Jahren Anstrengungen unternommen, die Transaktionskostentheorie zu einer „New Economics of Organization" auszubauen. (Vgl. *Williamson* 1981 a; 1985; 1990) Dieser Ansatz ist interdisziplinär (Ökonomik, Organisationstheorie, Recht) ausgelegt und versucht, im Gegensatz zur Standardökonomik, der Tatsache Rechnung zu tragen, daß die Form der Organisation einen entscheidenden Einfluß auf die wirtschaftliche Leistung hat. Die Erforschung der Bedeutung moralischer Kommunikation für die Durchführung und Sicherung von Transaktionen in der Unternehmung und zwischen Unternehmen steht auf der Forschungsagenda der Neuen Organisationsökonomik. (Vgl. *Williamson* 1985: 44[3], 268 ff., 405,; 1990: 177) Die Erträge dieser Diskussion sind bisher eher bescheiden, aber die Charakterisierung der Firma als System simultaner Markt- und Organisationssteuerung zur Allokation von Gütern und Koordination und Überwachung von Faktorleistungen könnte ein theorietechnischer Vorteil bei der Analyse unternehmensethischer Fragen sein. Sie erlaubt die Annahme der simultanen Relevanz von Ökonomie und Ethik in einem Ereignis. Da dieser Ansatz auch über ordnungs- und wirtschaftspolitische Ambitionen verfügt (vgl. *Williamson* 1985: 365 ff.; 1990: 200[3]), spricht einiges dafür, mit der so aufgespannten Theoriearchitektur die Erklärung der referierten unternehmensethischen Empirie zu versuchen. Ich werde so verfahren, daß ich zunächst den ordnungspolitischen, sodann den transaktionsökonomischen und schließlich den organisationstheoretischen Aspekt unter einem eher explorativen Gesichtspunkt diskutiere. Der Interferenzpunkt wird dabei die vertragstheoretische Interpretation von Transaktionen sein. Die kultur- und gesellschaftstheoretische Dimension und den Gesichtspunkt unternehmerischer „social responsibility" blende ich aus Gründen der Problemkonzentration aus. (Vgl. *Wieland* 1993 a; 1993 b)

Wenn man sich nach empirischen Korrelaten für eine *ordnungspolitische* Erklärung moralischer Kommunikation in der Wirtschaft umsieht, dann sollte dem Rechtssystem einige Bedeutung zukommen. Es war ja nicht allein (und nicht einmal in erster Linie) die moralische Empörung über Insider-Handel, Bestechung, Betrug und so weiter, die der Diskussion über die Ethik der Wirtschaft in den U.S.A. einigen Auftrieb gegeben haben, sondern vor allem deren juristisch herbeigeführte finanzielle Konsequenzen, die in einigen Fällen den betroffenen Unternehmen die Kapitalausstattung entzogen. Die seit dem 1. November 1991 gültigen „Federal Sentencing Guidelines" mit dem Schwerpunkt „Chapter Eight: Sentencing of Organizations" unterstreichen die Bedeutung ordnungspolitischer Maßnahmen. Im Kern laufen diese neuen Richtlinien zur Rechtsprechung auf einen Zuwachs der finanziellen Verantwortlichkeit von Organisationen, also auch und vor allem von Unternehmen, für ungesetzliches Handeln ihrer Mitglieder hinaus. Darunter fallen eine ganze Reihe von Vergehen, die eine dominante Rolle in der ethischen Diskussion spielen: Betrug, Bestechung, Umweltzerstörung. Die Strafen können bis zu mehreren hundert Millionen Dollar betragen, wenn das

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Unternehmen nicht nachweisen kann, daß es gebührende Anstrengungen unternommen hat, gesetzeskonformes Handeln seiner Akteure sicherzustellen. Unter gebührender Anstrengung wird dabei in den „Sentencing Guidelines" neben der Dokumentation einer wertegebundenen Unternehmenspolitik der Nachweis von firmeninternen Standards und Verfahren zur Umsetzung und Evaluierung dieser Politik verstanden. Das reicht von der Vorsorge, daß bekanntermaßen skrupellose Mitarbeiter nicht mit moralisch sensiblen Aufgaben betraut werden, über das Mitarbeitertraining zu Wertefragen bis hin zur Einführung eines Audit- und Frühwarnsystems. (Vgl. *Wieland* 1993a) Für den Nachweis eines installierten Programms dieser Art kann ein Richter auf eine Verringerung der Geldstrafe zwischen 20% und 60% erkennen. Wenn das Unternehmen zusätzlich den Fall selbst aufdeckt, kann die Reduktion 80% bis 95% betragen. Es ist nicht schwer zu sehen, daß ethische Unternehmensrichtlinien, Ethikkomitees, Ethikabteilungen und so weiter als gebührende Anstrengungen im Sinne dieser Richtlinien interpretiert werden können. Damit existiert ein erheblicher ordnungspolitischer Anreiz für Firmen, den moralischen Aspekt ihrer Transaktionen zu kontrollieren und zu routinieren. Man sollte sehen, daß die Richtlinien Vorkehrungen zur Einhaltung von Gesetzen verlangen, nicht die Internalisierung weitergehender Wertemuster. Die Erfahrung zeigt aber, daß ohne ein umfassendes unternehmensethisches Programm die Erfüllung rechtlicher Anforderungen in relevanten Fällen nicht sichergestellt werden kann. Moralische Kommunikation dient hier der Durchsetzung rechtlich codierter Sachverhalte. Dies erklärt zum Teil die bereits erwähnte Mischung wirtschaftlicher, politischer und moralischer Werte in den „codes of conduct". Im Abschnitt 3 werde ich diese „Instrumentalisierung" von Moral für Recht diskutieren. Ordnungspolitisch bleibt nur noch die Bemerkung, daß die „Sentencing Guidelines" nicht die Existenz von Ethikprogrammen (die schon weitaus länger existieren) erklären können, sondern vermutlich nur das rapide Wachstum seiner organisationellen Formen. Mit anderen Worten: Wir müssen uns nach weiteren Erklärungen umsehen.

Die Her- und Sicherstellung von Kooperationschancen bildet den Ansatzpunkt für eine *ökonomische* Erklärung der Unternehmensethik, weil Kooperation konstitutiv anfällig für „moral hazard" ist.[2] Der Kern des Kooperationsproblems liegt darin, daß bei den Beteiligten über die Beteiligten ein Wissen vorliegen muß, das es allen erlaubt, sich eine zufriedenstellende Vorstellung über den Verlauf der Kooperation zu machen. (Vgl. *North* 1990: 15) Wir können also formulieren, daß ein „code of ethics" jene Mindestmenge an Wissen über die moralischen Standards eines Akteurs offenlegen sollte, die notwendig ist, der unvollständigen Informationslage und der Unsicherheit seiner Kooperationspartner hinreichend zu begegnen. Unter der Bedingung beschränkter Rationalität geht das nur wirk-

---

2 Angelsächsich nüchtern definiert *Knight* (1965: 254) das Problem der Form Unternehmung als „the protection . . . of members and adherents against each other's predatory propensities."

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

sam, wenn der „Code" als Selbstbindung und Selbstverpflichtung formuliert ist.[3] Die den „Code" umsetzenden Ethikprogramme haben dann die Funktion, den Opportunismusvorbehalt, unter dem der Verhaltenskodex nach wie vor steht, weiter abzubauen. Die Umwandlung des Kodex in ein „living document" durch die Investition von Ressourcen und durch rechtliche Codierung signalisieren, daß die moralischen Kommunikationsangebote ernst gemeint sind. Nur durch den Einbau in die Instrumente der Unternehmensführung werden Ethikprogramme zu „credible commitments", zu ex ante angebotenen Sicherungen gegen ex post möglichen eigenen Opportunismus.

Ich folge mit dieser Argumentation den beiden grundlegenden Verhaltensannahmen, beschränkte Rationalität und Opportunismus, die Williamson für die Transaktionskostenökonomik vorgeschlagen hat. Nimmt man die Transaktion als Basiseinheit der Untersuchung, dann sollte diese so organisiert werden, daß möglichst viele transaktionsrelevante Informationen verfügbar sind und Opportunismus weitestgehend ausgeschaltet ist.[4] Zieht man in das Kalkül, daß die Firma als ein Netzwerk von Verträgen definierbar ist, das sowohl ex ante als auch ex post Transaktionskosten generiert, dann zeigt sich, warum Fragen der Moral in dieser Welt existieren und warum sie ökonomisch relevante Fragen sind. Sowohl der Schutz vor opportunistischem Verhalten als auch dessen Konsequenzen generieren Kosten. Würden sich alle ökonomischen Akteure an das Hobbessche Verdikt halten, daß Moral sich vor allem im Einhalten von Vertragsversprechungen zeigt (vgl. *Hobbes* 1651/1914: 74), dann würde das in der Tat den denkbar wirksamsten und kostengünstigsten Schutz gegen Opportunismus von Vertragspartnern bilden. Allerdings könnten unter solchen Voraussetzungen auch alle Transaktionen über den Markt abgewickelt werden, eine Notwendigkeit für die Organisationsform „Firma" existierte nicht. Da die Eintrittswahrscheinlichkeit dieses Ereignisses, so lehrt die Erfahrung, bedauerlicherweise aber nie den Wert 1 annimmt, sind kostenwirksame Vorkehrungen, also Organisation, erforderlich. Dieser Sachverhalt erlaubt es, in einer zweiten, nun ökonomischen Näherung, die Implementierung von Ethikprogrammen in der Unternehmung als Erweiterung und Justierung der „governance structure" einer Wirtschaftsorganisation zu verstehen, für die es gute Gründe gibt.

Das systematische Auftreten von Opportunismus ist eine der zwei zentralen Verhaltensannahmen der Organisationsökonomik. Dies nicht in dem Sinne einer

---

[3] Als Beispiel für eindeutige Formulierungen sei die Präambel der „Standards of Business Conduct" der Firma Northrop angeführt: „We are each personally accountable for high standards of behavior, including honesty and fairness in all aspects of our work. We strive to fulfill our commitments as law-abiding citizens and conscientious employees. We will be responsible and reliable, truthful and accurate, and diligent in the effective use of company and customer resources."

[4] Vgl. *Williamson* (1981 a: 1546) und (1985: 32): „The organizational imperative that emerges in such circumstances is this: Organize transactions so as to economize on bounded rationality while simultaneously safeguarding them against the hazards of opportunism."

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

anthropologischen Konstante, sondern als Behauptung über die Welt: Die Erfahrung zeigt, daß nicht alle Vertragspartner ihre Versprechen halten, sondern daß einige die Ausbeutung von Vertragsbeziehungen präferieren. Damit ist auch gesagt, daß der Versuch, ex ante die „good guys" von den „bad guys" zu unterscheiden, nicht vollständig möglich ist und Kosten dafür anfallen.[5] Diese Verhaltensannahme unterscheidet sich prinzipiell von solchen, die mit Kohlberg annehmen, individuelle moralische Kompetenz sei ein Hauptfaktor gelingender Unternehmensethik. Organisationsökonomisch ist Opportunismus kein Fall von Inkompetenz, sondern der moralische „worst case": Achtung! Mit kalkulierender Hinterhältigkeit muß gerechnet werden. Damit ist Tugendhaftigkeit nicht ausgeschlossen, sondern deren Vorliegen verbessert die Transaktionsbedingungen. Aber binden, so wäre das Argument, kann sich eine empirisch robuste Unternehmensethik an diesen kontingenten Sachverhalt in komplexen Wirtschaftsbeziehungen nicht. Pointiert formuliert: Die Unternehmensethik der Organisationsökonomik rechnet nicht mit Akteursmoral, sondern setzt auf Organisationsmoral. Natürlich gibt es letztere ohne erstere letztlich nicht, aber solche Aussagen nehmen nur allzuleicht die Differenzierungsgewinne zurück, die man mit der Unterscheidung von individuellen und kollektiven moralischen Akteuren hinsichtlich einer Erklärung der Existenz institutionalisierter moralischer Kommunikation in der Wirtschaft erzielen kann.

Wenn wir annehmen, daß die Firma ein Netzwerk verschiedener Vertragsformen ist, und zunächst einmal intrafirm- und interfirm-Verträge unterscheiden (vgl. *Alchian / Woodward* 1987: 110), dann fallen die in Abschnitt I erwähnten Fälle von Bestechlichkeit einzelner Teammitglieder, die private Nutzung von Firmeneigentum und opportunistische Leistungserbringung (Quantität und Qualität der Arbeitsleistung) unter die erste Rubrik.[6] Unter die zweite Kategorie rubrizieren wir alle Fälle von Lieferantenopportunismus (Produktqualität, Liefertermine, Garantieleistungen) und Opportunismus der Firma gegenüber Kunden (unfaire Preisabrechnung, Bestechlichkeit).

Die antiopportunistische Funktion von Ethikprogrammen besteht grundlegend darin, diese in der Praxis wohlbekannten „grey areas" offen zu benennen und potentiellen Kooperationspartnern zu versichern, solche Praktiken von Mitgliedern weder zu erwarten noch zu dulden und auf ihre Vermeidung hinzuarbeiten. Es wäre ein fataler Irrtum anzunehmen, diese Selbstverpflichtungen seien allenfalls guter Stoff für clevere PR-Arbeit. Die Wahrheit ist, daß solche Erklärungen Handlungserwartungen an die Organisation und ihre Mitglieder generieren, von denen unklar ist, ob man sie, selbst bei bestem Willen, immer wird erfüllen

---

[5] *Alchian / Woodward* (1988: 67) weisen darauf hin, daß selbst bei bestem Willen zum Fairplay diese Kosten anfallen würden, da unterschiedliche Problemperzeptionen nicht ausgeschlossen werden können.

[6] Vgl. *Wieland* (1993a: 45 f.) Dort findet sich auch die Auswertung der „Open Line" bei Northrop, die zu ähnlichen Schwerpunkten führt.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

können. Erwartungsenttäuschung wirkt dann kontraintentional. Moralische Kommunikation in der Wirtschaft ist risikoreich, weil sowohl die Erfüllung als auch die Dementierung von Erwartungen kostenpflichtig ist. Die Schaffung von Erwartungssicherheit und die Absorbierung des daraus folgenden Risikos erzwingen die Ausgestaltung der verbalen Verpflichtung auf einen Verhaltensstandard zu einer glaubwürdigen Versicherung. Ein Weg, dies zu tun, ist die Institutionalisierung von Ethikprogrammen, die (wie die hier beschriebenen) Kontrollelemente enthalten. Ich werde später zeigen, daß unter diesen Bedingungen selbst Erwartungsenttäuschungen zu Reputationsgewinnen führen können.

Ein zweiter Grund, den die Organisationsökonomik zur Erklärung von Organisationsmoral nahelegt, ist Faktorspezifität. Die Wirksamkeit von Opportunismus ist nicht nur eine Folge beschränkter Rationalität, sondern auch determiniert durch die Häufigkeit von Transaktionen und durch den Grad der Transaktionsspezifität von Investitionen.(Vgl. *Williamson* 1985: 31, 56) Hohe Faktorspezifität macht Transaktionen für Opportunismus anfällig. Äquivalent und auf unser Thema hin formuliert: Dort sind die Erträge aus antiopportunistischen Programmen hoch. Für diese Deduktion gibt es empirische Korrelate. Als Beispiel soll der Markt für militärische Güter herangezogen werden. Hier besteht zwischen dem Anbieter und dem Abnehmer eine hohe wechselseitige Abhängigkeit, die sich an entsprechenden Investitionen festmacht. So erfordert die Entwicklung von Waffensystemen und auch deren Vertrieb eine enge und langfristig angelegte Zusammenarbeit zwischen verschiedenen Anbietern und den Anbietern und dem Abnehmer. Die Investitionen sind auf allen Seiten hoch. Zwar kann der Abnehmer unter der Bedingung internationaler Konkurrenz im Prinzip zwischen verschiedenen Anbietern wählen, aber diese Macht ist durch militärische und politische Interessen und auch durch die technologische Abhängigkeit von einmal eingegangenen Kooperationen stark eingeschränkt. Für die Anbieterseite existiert im Prinzip ein (legaler und illegaler) nationaler und internationaler Markt, der aber durch die eben erwähnten Abnehmerinteressen eingeschränkt wird. Der Anreiz, Kosten während der Entwicklungsphase von Systemen (die der Anbieter nach Vertragsabschluß organisiert) opportunistisch zu deklarieren und Skalenerträge für einmal entwickelte Waffensysteme zu maximieren, verletzt die Interessen des Auftraggebers, was von diesem, bei Aufdeckung, negativ sanktioniert werden wird. Die Abhängigkeit in diesem Markt mag symmetrisch sein, nicht aber die Möglichkeit zu ex post Opportunismus. Organisationsökonomisch kann man in diesem Fall zunächst einmal prognostizieren: „where asset specificity is great, buyer and seller will make special efforts to design an exchange relation that has good continuity properties.“ *(Williamson* 1981: 1546)

Dies ist in unserem Beispiel und zugeschnitten auf unser Problem der Fall: Die Unternehmen der Rüstungsindustrie in den U.S.A. waren und sind eine der treibenden Kräfte in dem hier diskutierten Segment der Unternehmensethik. Die Reaktion des Pentagon auf Bestechungsskandale bei der Vergabe von Entwick-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

lungsprojekten oder internationalen Aufträgen und falsche Leistungsabrechnungen waren und sind die konkreten Anlässe. Die Pentagon Defense Logistics Agency schließt Lieferanten, die in Betrugs- und Bestechungsaffären verwickelt sind, konsequent von der Vergabe neuer Aufträge solange aus, bis die betroffene Firma (in dem nun zitierten Fall aus 1992: General Electrics) „has tightened its internal control and ethics programms . . .". (New York Times, 3.6.1992: D1) Fast alle Lieferanten des Pentagon waren schon Objekt solcher Maßnahmen. Entscheidender Aspekt für unsere Diskussion ist, daß sie wegen der diskutierten bilateralen Abhängigkeit nicht auf den sofortigen Abbruch, sondern auf die Kostenspezifikation von Zukunftstransaktionen (Verlust an Reputation, Verlust von Märkten) zielen. Die New York Times bemerkt daher zum temporären Charakter des Abbruchs der Geschäftsbeziehungen richtig: „But the action clearly demonstrated the Pentagon's willingness to use publicity and the threat of financial harm in the long run to put pressure on the company to keep a tighter rein on employes and rectify apparent ethical lapses . . .". (ebenda)

Der Druck von Publizität ist wirksamer, wenn er sich auf öffentliche Selbstbindungen (code of conduct) bezieht, die als glaubwürdige Zusicherungen der Anbieter im Hinblick auf hohe transaktionsspezifische Investitionen gedacht sind. Glaubwürdigkeit wird hier aber nicht durch Bedauern ausdrückende Stellungnahmen und Zusicherung von Besserung erreicht, sondern durch den Kosten generierenden Einbau von den eigenen Opportunismus aufdeckenden Mechanismen in der „governance structure" des Unternehmens. Boeing nennt sein Ethikprogramm nicht ohne Grund ein Programm zur „self-governance". Das Abstellen auf Reputationsgewinne, die sich in jedem Fall (also sowohl bei der Aufdeckung und Verarbeitung von Opportunismus als auch bei seiner Vermeidung) bei dieser Strategie einstellen, ist dann am effizientesten, wenn die deklarierten Ansprüche eindeutig und leicht zu überwachen sind. (Vgl. *Kreps* 1990: 116) Dies erklärt, warum sich 45 Unternehmen, die gemeinsam mehr als 50% des Umsatzes auf dem Markt für militärische Güter tätigen, zur bereits erwähnten „Defense Industry Initiative on Business Ethics and Conduct" (DII) zusammengeschlossen haben. Mitglied der DII kann werden, wer die Umsetzung ethischer Unternehmensrichtlinien aktiv betreibt und dazu mindestens ein Ethikkomitee einrichtet. Der Jahresbericht dieser Organistion für 1990 weist aus, daß alle beteiligten Unternehmen zusätzlich über eine „hot line" und einen Ombudsmann verfügen. Fassen wir zusammen: Die hier diskutierten unternehmens- und branchenspezifischen Ethikprogramme sind ex ante angebotene Wissensbestände und Sicherheiten der Anbieter und Lieferanten, die die angestrebte Integrität der Transaktionen glaubwürdig signalisieren sollen. In der Terminologie von Williamson sind sie „safeguards" vom Typ II (Signalisieren einer Präferenz für außergerichtliche und gütliche Lösungen von auftretenden Konflikten) und Typ III (Signalisieren einer Präferenz für Kontinuität). (Vgl. *Williamson* 1985: 34) Die Annahme von einer Präferenz für Kontinuität ist die ökonomische Übersetzung des organisationstheoretischen Begriffs „Firmenidentität". Daß Identität unternehmenspolitisch wichtig ist, ist

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

aus dieser Sicht entweder die Konsequenz der Transformation anonymer Marktbeziehungen in langfristig angelegte Vertragsbeziehungen oder wird von Anfang an durch absehbar hohe transaktionsspezifische Investitionen induziert. In beiden Fällen tritt Identität der Vertragspartner an die Stelle ihrer Anonymität.[7]

Es sollte daher möglich sein, die Implementierung organisationell abgestützter moralischer Kommunikation als eine Funktion des Verhältnisses von Reputationsgewinnen (vermiedene Kosten aus aufgedecktem Fehlverhalten; Gewinne aus zusätzlichen Kooperationschancen) zu den Opportunitätskosten fairen Verhaltens (entgangene Gewinne aus Opportunismus) zu verstehen. Überwiegen die Gewinne die Kosten und ist die Faktorspezifität signifikant (motivational: es gibt eine Präferenz für kontinuierliche Zusammenarbeit), dann gibt es eine Anreizkonstellation, die zur Emergenz von „institutional and personal trust relations" *(Williamson* 1981b: 240) führen kann. Dies gilt auch, wenn sowohl Gewinne als auch Kosten absolut signifikant sind, ihre Relation daher indifferent ist. Wenn opportunistisches Verhalten schnell entdeckt werden kann, das Wissen über dieses Fehlverhalten einen großen Ausdehnungsraum besitzt und ein Ausweichen auf andere Märkte nicht möglich ist (asset specificity), dann erhöhen sich mit der Zahl kontinuierlicher Transaktionen (frequency) die Opportunitätskosten von Opportunismus. (Vgl. *Milgrom / Roberts* 1988: 449) In Gesellschaften, deren Wirtschaft eher durch Organisation als durch Spotmärkte charakterisiert ist, gewinnt daher der Faktor Moral, die Reputation als „asset", in der Wirtschaft an Bedeutung. Umgekehrt, umgekehrt. Die Tatsache, daß moderne Gesellschaften Organisationsgesellschaften[8] sind, wäre demnach eine bedeutende Triebfeder der Renaissance von „Business Ethics".[9] Eine ökonomische Theorie der Organisation ist mit ihrer Theoriearchitektur auf diese Entwicklung eingestellt.

Das führt uns zum *organisationstheoretischen* Aspekt unseres Themas. Ich bin bisher implizit davon ausgegangen, daß die Firma die Organisationsform eines Teams von Faktoranbietern ist, die heterogene Präferenzen und potentiell konfligierende Nutzenfunktionen haben. Kooperation ist unter diesen Bedingungen das Ergebnis eines Anreiz- und Überwachungssystems, welches die Konsistenz der Handlungen erzeugt. Präferenzen sind exogen und stabil (pro Opportu-

---

[7] Vgl. *Williamson* (1985: 62): Darin ist gesellschaftstheoretisch impliziert, daß moderne Wirtschaften keinesweg nur durch anonyme Marktbeziehungen (Spot-Märkte) charakterisiert sind. Unter der Bedingung von „Markets and Hierarchies" hängt es vom Dimensionenmix der Transaktion ab, ob Anonymität oder Identität effizient ist. Diese Einsicht hat Konsequenzen für Wirtschaftsethiken, die mit Hinweis auf Arbeitsteilung und Anonymität den systematischen Ort der Ethik in der staatlichen Rahmenordnung lokalisieren.

[8] Vgl. für diese Beobachtung schon früh Knight (1941/1982: 207 f., 208[4]). In die gleiche Richtung gehen Beobachtungen der Organisationssoziologie (*Perrow* 1989) und der Politikwissenschaften (*Mayntz* 1992).

[9] Vgl. zu dem hier entwickelten Argument samt Schlußfolgerung *Milgrom / Roberts* (1988: 449): „This suggests that a society of long-lived, formal organizations may be especially effective in using reputation mechanisms. Conversely, the reputation mechanism cannot operate effectively in fluid, impersonal market settings."

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

nismus) in diesem Modell. Daraus folgt, daß Ethikprogramme nicht auf die Homogenisierung der Motivationsstruktur der Teammitglieder zielen, die dann Kooperation erzeugen würde. Die empirische Untersuchung lieferte für genau die entgegengesetzte Auffassung Belegmaterial. Große Unternehmen, die international operieren und ihre Mitglieder in pluralistischen und multikulturellen Gesellschaften akquirieren, stehen vor dem Problem, daß individuell und kulturell divergierende und nicht kompatible Wertemuster der Mitglieder die Existenz von Opportunismus in der Organisation begünstigen. Zwar sind Bestechung und Betrug in allen Ländern der Erde auf die eine oder andere Weise strafbar, aber ihre normative Bewertung ist sehr unterschiedlich. Wenn daher ein Team über die moralischen Standards seiner Mitglieder kein hinlänglich sicheres Wissen erreichen kann, gleichzeitig aber möglicherweise mit weitreichenden Konsequenzen nichterwünschter Standards konfrontiert werden kann, dann liegt ein Ausweg aus dieser Situation in Maßnahmen, die auf eine lokale Homogenisierung moralischer Werte zielen. Das heißt praktisch, daß der kollektive Akteur gegenüber seinen individuellen Akteuren klarstellt, was die kollektiven Werte (also die der Form „Unternehmung", unabhängig von ihren Mitgliedern) sind und deren Durchsetzung und Einhaltung betreibt. Die Bedeutung dieses Aspektes wächst mit der Größe und Komplexität der Organisation und mit der Existenz von internen Märkten.[10] Daran kann eine organisationelle Deutung institutionalisierter Ethikprogramme anschließen. Sie sind dann als ein auf Kontextsteuerung abstellender Mechanismus der Selbstbeobachtung und Selbstkontrolle (self-governance) eines kollektiven Akteurs zu verstehen, der intern auf die Herstellung einer homogenen und integrativen Firmenkultur und Identität und extern auf die Herstellung von Vertrauenswürdigkeit, Akzeptanz und Autorität zielt. Wertemanagement und Moralcontrolling sind ein Mittel zur Kontextsteuerung und Kontrolle komplexer Organisationen, sowohl hinsichtlich ihrer internen als auch externen Relationen durch Homogenisierung der organisationellen Verhaltensstandards. Dies geschieht erstens dadurch, daß die moralischen Lernprozesse des Unternehmens gefördert werden, daß zweitens die moralische Kommunikation in der Firma und zwischen der Firma und ihrer Umwelt zum Gegenstand gezielter Unternehmenspolitik wird und daß drittens Wertefragen im Unternehmen routinemäßig Aufmerksamkeit finden. In allen von mir durchgeführten Experteninterviews wurde deutlich, daß ein „code of conduct" und seine Umsetzungsinstrumente auch in diesem Sinne verstanden werden.

Muß eine Organisationsökonomik daraus den Schluß ziehen, daß, zumindest in diesem Fall, endogene und instabile Präferenzen als Modellannahmen ange-

---

[10] Die Aufgliederung und anschließende Vernetzung großer Unternehmen in rechtlich, aber nicht wirtschaftlich unabhängige „interne Märkte" läuft auf die Wiedereinführung von Markt in Organisation hinaus. Dort gelten dann als Verhaltensnorm Kooperation und Wettbewerb. Die Frage erhebt sich, was es unter diesen Bedingungen noch heißt, für das Unternehmen X oder Y tätig zu sein. Eine Möglichkeit, dieses Identitätsproblem zu lösen, sind „codes of conduct".

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

messen sind, wenn empirische Relevanz nicht aufgegeben werden soll? Dagegen steht die Forderung an jede Theorie, den Komplexitätszuwachs einer Fragestellung so weit zu begrenzen, daß er handhabbar bleibt und ihm ein entsprechender Zuwachs an Erklärung entspricht. Dies gilt vor allem für Forschungsansätze, die die Notwendigkeit interdisziplinärer Arbeit betonen. Wenn man Interdisziplinarität nicht als Schaffung einer ganzheitlichen Metatheorie ansetzt, dann läuft sie meines Erachtens auf die Forderung hinaus, die Beobachtungen perspektivisch anders eingestellter Forschungsansätze zunächst einmal als Erklärungsaufforderung für das eigene Modell zu verstehen. Darin impliziert ist die Akzeptanz differenter Perspektiven als erwünschter Zugewinn an Problembeschreibung und sich daran anschließender Handlungsoptionen. Kurz und gut, ich plädiere für ein Übersetzungsprogramm, das zwei Fragen beantworten muß: Ist es auf der Basis des methodologischen Individualismus (heterogene Präferenzen, potentiell konfligierende Nutzenfunktionen) möglich, den Aspekt angestrebter motivationaler (moralischer) Homogenität der Ethikprogramme zu erklären? Wie ist es möglich, zugleich von individuellem Opportunismus und angestrebten kollektiven homogenen moralischen Präferenzen auszugehen? Anders formuliert: Inwiefern kann man sinnvoll von der Existenz eines kollektiven moralischen Akteurs ausgehen?

Bekanntlich hat Coase den Weg zu der Einsicht geöffnet, daß die Emergenz der Firma die Umstellung wirtschaftlicher Allokation und Koordination von Spotmärkten auf relativ langfristige vertragliche Arrangements bedeutet. Solche Arrangements schränken die gegenwärtigen und zukünftigen Handlungsmöglichkeiten der involvierten Tauschpartner ein. Das ist ihr Sinn, und nur dadurch wird Teamarbeit möglich. Die Unternehmung als „Nexus von Verträgen" konstituiert sich durch die Einschränkung der Handlungsmöglichkeiten ihrer Mitglieder. Hobbes Einsicht, daß konstitutionelle Akte für das involvierte Individuum die Aufgabe „des Rechts auf alles" bedeuten, greift auch hier. Vor den bereits erwähnten intrafirm und interfirm Verträgen steht ein konstitutioneller Akt, der die Firma als kollektiven Akteur überhaupt erst bildet. (Vgl. *Alchian / Woodward* 1987: 110)

Diese Argumentation erlaubt es uns, ethische Unternehmensstandards als konstitutiv für die Etablierung der Unternehmung als kollektiven moralischen Akteur zu verstehen, und zwar sowohl gegenüber den Teammitgliedern als auch ihrer Umwelt. Wir sehen darin weiterhin den Versuch von Firmen, zwischen gesellschaftlichen (Rahmenordnung) und individuellen Verträgen ein Kontinuum von Selbstbindungsmechanismen zu implementieren. Beide Überlegungen möchte ich nun weiter erläutern.

Mit dieser theoretischen Disposition können wir Unternehmen als kollektive wirtschaftliche und moralische Akteure verstehen, ohne die Logik des methodologischen Individualismus zu verlassen. Sie sind kollektive Akteure in dem Sinne, daß sie sich als solche konstituieren und dann in organisierter Weise Entscheidungen herbei- und durchführen. Darin ist eingeschlossen, daß die einzelnen Mitglieder eigene Interessen mit ihrer Unternehmenszugehörigkeit verfolgen. Sie können

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

dies aber nur, insofern sie die Unternehmensverfassung und die kodifizierten (einschließlich der impliziten) Verhaltensstandards als Einschränkung ihrer Handlungsfreiheit akzeptieren.[11] Von einer Homogenisierung moralischer Präferenzen kann also nur gesprochen werden, insofern zwischen dem kollektiven und dem individuellen Akteur eine Differenz liegt, die dem ersteren eine distinkte Existenzform und dem letzteren Entscheidungsspielräume gibt. Anderenfalls wäre die Existenz von Opportunismus unerklärlich. Geht man davon aus, überrascht einen die Tatsache nicht, daß weder Verhaltenskodex noch ausgefeilte Ethikprogramme Betrug und Bestechung (so geschehen sowohl bei Boeing als auch bei Northrop) ausschließen können. Moralität ist kein herrschender Zustand, sondern ein sein Ziel nie erreichender Prozeß. Was man daher erklären muß, sind nicht allein die Bemühungen zur Homogenisierung der Präferenzen eines kollektiven Akteurs, sondern simultan die nichtkontingente Existenz von Opportunismus.

Ethikprogramme sind ein Akt der Selbstbindung aller Mitglieder des Teams, der heterogene moralische Präferenzen nicht ausschließt, sondern voraussetzt. Warum individuelle Akteure dieser Bindung durch Zustimmung zum Arbeitsvertrag zustimmen, kann sehr verschiedene Beweggründe haben. Das Spektrum reicht hier von der moralischen Persönlichkeit über das Streben nach Achtung der eigenen Person bis hin zur Kalkulation des Zukunftsnutzens einer Investition in Reputation, Ehrlichkeit, Fairneß und so weiter.[12] Das Problem scheint mir eher darin zu liegen, daß die Vertragspartner einem ziemlich vagen Dokument zustimmen und ex ante dem Prinzipal das Recht einräumen, ex post die Konsequenzen (Kosten) daraus zu spezifizieren, die leicht aus „unforeseen contingencies" (Kreps) resultieren können. Daß der Prinzipal in der Regel auch gegenüber seinen Agenten auf Kontinuität und Reputation achten muß, ist eine der Sicherungen gegen dessen mögliche Willkür. (Vgl. *Kreps* 1990: 113)

„Codes of ethics" sind Handlungsbeschränkungen, die organisationsintern auf eine Senkung der Preise für moralisches Verhalten[13] und auf eine Stimulierung der Nachfrage durch die Mitglieder zielen. Denn auch für diese gilt, daß moralische Kommunikation in der Wirtschaft risikoreich ist. Sie zielen weiterhin darauf,

---

11 Mit dieser Argumentation folge ich *Vanberg* (1992).

12 Auch die spieltheoretisch ausgelegte Teamtheorie hat die Tragweite dieser Diskussion registriert. So plädiert *Radner* (1992: 410) dafür, Prozesse der Normstandarisierung (also die Überprüfung der Annahme identischer Präferenzen) als Forschungsfrage in die Agenda aufzunehmen, selbst angesichts der Tatsache, daß damit der sichere Boden spieltheoretischer Ökonomik in Frage gestellt werden könnte. Dahinter steht die Einsicht, daß Principal-Agent-Beziehungen komplexer sind, als die Beiträge der in dieser Hinsicht „somewhat disappointing" Spieltheorie nahelegen.

13 Dieser Gesichtspunkt ist von großer Bedeutung. Wo dies nicht geschieht, können Moral präferierende Mitglieder in die Positition eines „whistleblowers" getrieben werden. An ihnen kann man die möglichen Preise für nichtorganisierte moralische Kommunikation ablesen. Nach einer amerikanischen Studie über 87 „whistleblowers" verloren 17 % ihr Haus, 8 % mußten Bankrott anmelden, 15 % wurden von ihren Ehepartnern geschieden und 10 % versuchten Selbstmord. Vgl. Financial Times, January 12, 1994: 7.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

durch interne wechselseitige Verpflichtung und externe Selbstverpflichtung (auf Reputation als „asset" der Organisation) stabile Handlungserwartungen dadurch zu schaffen, daß man sich selbst festlegt und damit Anreize für andere schafft, sich auch festzulegen. Dahinter steht die Absicht, „to invest in changing the behavior of the other in the direction of inducing more cooperation". *(Buchanan* 1991: 185; vgl. *Kreps* 1990; *Priddat* 1993) Das entscheidende Resultat dieser konstitutionellen Diskussion organisationstheoretischer Aspekte der Unternehmensethik ist dieses: Die begründend gemeinte These, Unternehmensethik sei möglich, weil und solange Handlungsspielräume existierten, trifft nicht den Kern des Problems. Unternehmensethik zielt neben der Stabilisierung vorhandener vor allem auf die Schaffung zusätzlicher Handlungsspielräume (Transaktionsoptionen) für die und in der Unternehmung. Der Korridor, in dem Kooperationsgewinne abgeschöpft werden können, dehnt sich aus. In einer Welt moralischer Anomalie hingegen geht die Zahl der möglichen Transaktionen aller Akteure und damit die möglichen Kooperationsgewinne gegen Null.

3.

Und die Ethik? Ich habe versucht, dafür zu argumentieren, daß institutionalisierte Formen moralischer Kommunikation in der Wirtschaft der Sicherstellung der Funktion von Firmen zuarbeiteten, nämlich der Allokation von Ressourcen und der Koordination und Überwachung von Produktionsfaktoren durch Organisation. Auch wurde auf die Funktion der Moral hingewiesen, an der Durchsetzung von Recht mitzuwirken. Ich habe an anderer Stelle den Zusammenhang von funktionaler Differenzierung und Wirtschaftsethik diskutiert und dafür argumentiert, daß dies unter den Anwendungsbedingungen moderner Gesellschaften nicht anders sein kann. (Vgl. *Wieland* 1993b)

Die Frage allerdings bleibt: Läuft dieser Einbau von Moral in die Funktionserfordernisse der Subsysteme differenzierter Gesellschaften nicht auf eine Instrumentalisierung der Ethik hinaus? Ich werde diese Frage in drei Schritten so strukturieren, daß eine mir plausibel erscheinende Antwort möglich wird. Dieses Verfahren ist ebenfalls eher explorativ als explanativ gemeint.

Mein erstes Argument ist, daß moderne Unternehmensethik primär nicht auf individuelle Tugendhaftigkeit abstellen kann. Nicht, daß diese unwichtig wäre. Im Gegenteil: Am Schluß sind es immer Individuen, die moralisch handeln oder nicht moralisch handeln. Moralische Kompetenz ist wichtig, in bestimmten Situationen auch entscheidend. Aber ich habe versucht, Argumente dafür vorzutragen, warum eine empirisch robuste Unternehmensethik davon nicht ihren Ausgangspunkt nehmen sollte.

Daran schließt sich ein zweites Argument folgenden Inhalts an: Die Umstellung auf ordnungspolitische Steuerung kann zwar auftretende Defizite eines indivi-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

dualethischen Ansatzes heilen, damit liegt aber noch keine vollständige Beschreibung moderner Unternehmensethik vor. Firmenspezifische Normstandardisierungen verweisen darauf, daß der Differenzierung moderner Gesellschaften ein tief gestaffeltes System konstitutioneller Vertragssysteme entspricht, die füreinander funktionale Äquivalente sein können. Unternehmensspezifische Ethikprogramme können eine Antwort auf die moralische Dekomposition der Gesellschaft sein, die die möglichen Kooperationschancen und Kooperationsgewinne in der Wirtschaft negativ tangiert (und sie werden von manchen Unternehmen in den U.S.A. auch so verstanden). Darin steckt ein Zeitproblem: Mit Entscheidungen in der Wirtschaft kann nicht gewartet werden, bis die moralische Erosion der Gesellschaft gestoppt und ein neuer Konsens über Werte aufgebaut ist. Die Festschreibung von Spielregeln durch das Organisationssystem „Staat" oder durch das Organisationssystem „Firma" unterscheidet sich nicht am Grad ihrer Reichweite und Universalisierbarkeit, sondern daran, daß sie einmal politisch (rechtlich) und einmal ökonomisch codiert sind. Untersucht werden muß daher nicht nur, wie Spielregeln (Rahmenordnung) auf Spielzüge (Akteure) wirken (vgl. exemplarisch *Homann* 1993), sondern wie verschiedene Formen institutioneller Beschränkungen aufeinander und auf die Spielzüge wirken. Ökonomisch handelt es sich dabei um ein Effizienzproblem. Moderne Gesellschaften sind Organisationsgesellschaften, und das heißt, es gibt für prinzipiell alle Probleme mehr (und möglicherweise effizientere) Problemlöser als das Individuum und die politische Organisation „Staat".

Das bringt mich zu der dritten Überlegung: Unternehmensethik handelt von der Moral eines kollektiven Akteurs. Zwischen Individualethik und Ordnungsethik existierte demnach noch als dritte Form die Organisationsethik. Die soziologische Diskussion (vgl. *Geser* 1988) hat gute Argumente dafür vorgetragen, daß moderne Gesellschaften nicht nur ein Interesse daran haben, ihre Organisationen normativ einzubinden, sondern daß Organisationen wegen ihrer rational-strategischen Potentiale erhöhte Verantwortlichkeit, Erwartbarkeit und Fähigkeit zur Begründung und Realisierung moralischer Ansprüche zugerechnet werden. Spitz formuliert: Individuen dürfen moralisch versagen, Organisationen nicht. Dies könnte nicht nur erklären, warum speziell Unternehmen sich immer häufiger mit moralischen Ansprüchen konfrontiert sehen, sondern darin steckt auch ein Hinweis auf die Emergenz diverser Partikularethiken (Bioethik, Medizinethik, Sportethik, Wirtschaftsethik und so weiter). Deren gemeinsames Charakteristikum ist es, in einem lokalen Kontext[14] universalistisch begründete Ethiken anzuwenden. Das zwingt, den normativen und den deskriptiven Status ethischer Theorien zu unterscheiden. Die empirische Leistungsfähigkeit einer Ethik hängt von deren lokaler Kontextualisierung, ihrer Deskriptivität ab. (Vgl. *Vossenkuhl* 1993)

---

[14] Vgl. zum Begriff der „lokalen Gerechtigkeit" *Elster* (1992), dem ich in dieser Hinsicht weitgehend folge.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Diese drei Überlegungen führen mich zu dem von verschiedenen Autoren bereits vorgetragenen Argument, daß in lokalen Ethiken (also auch der Unternehmensethik) zwischen deren Begründung, Anwendung und Gültigkeit unterschieden werden muß. (Vgl. *Homann* 1993)

Begründung ist die autonome Aufgabe des Moralsystems und seiner Reflexionstheorien. Begründung zielt immer auf Universalisierung, und ihre Leistung für die Akteure der Gesellschaft besteht darin, zwischen begründbaren und nichtbegründeten Normen zu diskriminieren und erstere für lokale Diskurse zugriffsfähig zu halten. Begründungsdialoge und Instrumentalisierung schließen einander aus.

Die Anwendung moralischer Sätze erfolgt immer lokal und läuft auf die Balancierung differenter Entscheidungslogiken (in der Wirtschafts- und Unternehmensethik etwa: Ökonomie, Ethik, Technik, Recht) hinaus. Daraus folgt, daß Wirtschafts- und Unternehmensethik deskriptive Ethiken sind, die sich auf lokale Anwendungsdiskurse beziehen. Dabei ist prinzipiell völlig offen, ob die moralische Codierung eines Problems überhaupt etwas zu dessen (positiver, effizienterer) Lösung beiträgt. Die Erfahrung legt hier eine gesunde Skepsis nahe. Daß Moral Probleme der Wirtschaft löst, die durch wirtschaftliche oder etwa rechtliche Mittel nicht zu lösen sind, kann daher nicht die Denkvoraussetzung einer Wirtschafts- und Unternehmensethik sein, sondern der durch sie zu erbringende Nachweis. Verfügt sie aber über ein distinktes Problemlösungspotential und wird dies durch die Systeme und Akteure der Gesellschaft (also nicht nur der Wirtschaft) für ihre Funktionen und Ziele genutzt, dann ist genau das ein instrumenteller Akt. Eine lokale Anwendung von Ethik ohne Problemlösungspotential ist im besten Fall schlicht nutzlos. In auf Leistungsspezifität abstellenden funktional differenzierten Gesellschaften gilt aber auch der umgekehrte Sachverhalt: Die Moral nutzt („instrumentalisiert") die Leistungsfähigkeit der Wirtschaft zur Erreichung ihrer Ziele. Auch dagegen ist nichts einzuwenden, denn moralischer Kommunikation ohne Ressourcenausstattung kann nur in seltenen Fällen Wirksamkeit zugetraut werden.

Über die Gültigkeit moralischer Sätze entscheidet ihre lokale Anwendbarkeit.[15] Nichtanwendbare moralische Sätze gelten nicht, obwohl sie begründbar sein können. Der Grad der Instrumentalisierbarkeit entscheidet daher über die Gültigkeit. In den Anwendungsdiskursen funktional differenzierter Gesellschaften geht es daher immer um die wechselseitige Instrumentalisierung differenter Entscheidungslogiken. Daran kann es keine Kritik geben, es sei denn, man billigt morali-

[15] Vgl. für dieses Argument *Homann* (1993). Man kann der dort vorgetragenen Überlegung folgen, daß die Ergebnisse der Anwendungsdiskurse auf die Begründungsebene durchgreifen. Dann wäre Wirtschaftsethik in dieser Hinsicht ein Begründungsprogramm. Die Frage ist allerdings, ob damit nicht die Gewinne aus der Differenzierung von philosophischen und sozialwissenschaftlichen Diskursen verschenkt werden.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

scher Kommunikation a priori eine höhere Dignität zu als ökonomischer Kommunikation. Das ist in Begründungsdiskursen möglich, lokal umzusetzen ist es nicht.

Die Antwort auf die Frage nach dem instrumentellen Charakter einer Unternehmensethik lautet demnach: Das mit kritischen Intentionen vorgebrachte Instrumentalisierungsverbot beruht entweder auf der Verwechslung von Begründungs- und Anwendungsdiskursen oder auf der (wenig wahrscheinlichen) alteuropäischen Hoffnung, daß beide unter dem Primat der Ethik zusammenfallen.

Die Argumentation in diesem Aufsatz folgte daher auch einer anderen Überzeugung: Moral und Ökonomie sind in der Wirtschafts- und Unternehmensethik nur Einheit als und weil Differenz. Diese Paradoxie muß durchgehalten werden, wenn die jeweiligen Zwecke erreicht werden sollen. Der ökonomische Zweck (Senkung von Transaktionskosten, Sicherung von zukünftigen Kooperationschancen und so weiter) kann nur erreicht werden, wenn moralische Kommunikation als solche zugelassen wird. Der moralische Zweck (Zuweisung von Achtung / Mißachtung, Herstellung von Solidarität und so weiter) kann nur erreicht werden, wenn Ökonomie zählt. Abstrakter formuliert: Instrumentalisierung ist nur möglich, wenn Differenz (also Autonomie) akzeptiert wird. (Vgl. *Frank* 1987: 77)

In diesem Kontext erscheinen theoretisch mögliche Versuche, Ethikprogramme als PR-Kampagne und schlicht gewinnmaximierendes Managementinstrument zu begreifen, als für die Unternehmung gefährlicher Irrtum. Um was es geht, ist das risikoreiche Geschäft der Zulassung, Akzeptanz und Balancierung differenter Entscheidungslogiken in lokalen Situationen unter Entscheidungsdruck. Die Institutionalisierung moralischer Kommunikation in der Unternehmung sorgt hier für Entlastung und fördert Entscheidbarkeit. Aber auch das ist nur eine Version der nicht zu überschätzenden Tatsache, daß Sollen empirisch mit Können zu tun hat.

Literatur

Alchian, A. A. / *Woodward,* S., Reflections on the Theory of the Firm. In: Journal of Institutional and Theoretical Economics: 1987, 110-136.

— The Firm is Dead; Long Live the Firm — A Review of Oliver E. Williamson's The Economic Institutions of Capitalism. In: Journal of Economic Literature: 1988 (26), 65-79.

Buchanan, J. M., The Economics and the Ethics of Constitutional Order. Ann Arbor: 1991.

Center for Business Ethics at Bentley College, Are Corporations Institutionalizing Ethics?, In: Journal of Business Ethics: 1986 (5), 85-91.

— Instilling Ethical Values in Large Corporations. In: Journal of Business Ethics: 1992 (11), 863-867.

Elster, J., Local Justice. How Institutions Allocate Scarce Goods and Necessary Burdens. New York: 1992.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Frank, R., If Homo Economicus Could Choose His Own Utility Function, Would He Want One With a Conscience? In: American Economic Review: 1987 (77), 593-604.

Geser, H., Interorganisationelle Normkulturen. In: Haller, M. et al. (Hg), Kultur und Gesellschaft. Frankfurt a.M.: 1988.

Hobbes, Th., Leviathan. London: 1651/1914.

Homann, K., Wirtschaftsethik. Die Funktion der Moral in der modernen Wirtschaft. In: Wieland, J., 1993b.

Knight, F., Risk, Uncertainty and Profit. New York: 1965.

— Religion and Ethics in Modern Zivilisation. In: Freedom and Reform. Essays in Economics and Social Philosophy. Indianapolis: 1982.

Kreps, D., Corporate Culture and Economic Theory. In: Alt, J. / Shepsle, K. (eds.), Perspectives on Positive Political Econony. Cambridge: 1990.

Luhmann, N., Gesellschaftsstruktur und Semantik. Bd. 3. Frankfurt a.M: 1989.

Mayntz, R., Modernisierung und die Logik von interorganisatorischen Netzwerken. In: Journal für Sozialforschung 1992 (32, 1), 19-32.

Milgrom, P. / *Roberts,* J., Economic Theories of the Firm: past, present, and future. In: Canadian Journal of Economics: 1988 (XXI, 3), 444-458.

North, D. C., Institutions, Institutional Change and Economic Performance. Cambridge: 1990.

Perrow, Ch., Eine Gesellschaft von Organisationen. In: Journal für Sozialforschung 1989 (29, 1), 3-19.

Priddat, B. P., Ökonomische Moral und moralische Ökonomie. In: Geschichte und Gegenwart 1993 (2), 95-112.

Purcell, T. V. / *Weber,* J., Institutionalizing Corporate Ethics: A Case History. In: Special Study No 71 (The Presidents of the American Management Association, New York): 1979.

Radner, R., Hierarchy: The Economics of Managing. In: Journal of Economic Literature: 1992 (30), 1382-1415.

Vanberg, V. J., Organizations as Constitutional Systems. In: Constitutional Political Economy: 1992 (3, 2), 223-253.

Vossenkuhl, W., Normativität und Deskriptivität in der Ethik. In: Eckensberger, L. H. / Gähde, U. (Hg.), Ethische Norm und empirische Hypothese. Frankfurt a.M.: 1993.

Wieland, J., Formen der Institutionalisierung von Moral in amerikanischen Unternehmen. Bern / Stuttgart: 1993a.

— Die Ethik der Wirtschaft als Problem lokaler und konstitutioneller Gerechtigkeit. In: Ders. (Hg.), Wirtschaftsethik und Theorie der Gesellschaft. Frankfurt a.M.: 1993b.

Williamson, O. E., The modern corporations: Origins, evolution, attributes. In: Journal of Economic Literature: 1981a (19), 1537-1568.

— Transaction-Cost Economics: The Governance of Contractual Relations. In: Journal of Law and Economics: 1981b (22), 231-261.

— The Economic Institutions of Capitalism. New York: 1985.

— Organization Theory. From Chester Barnard to the Present and Beyond. Oxford: 1990.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Anhang

Schaubild 1

OUR MISSION

Northrop provides products and services which contribute to the defense and technological strength of our nation. Our goal is to be one of the top two firms in our chosen segments of the military/commercial aircraft and defense electronics markets. Our aim is to satisfy customer needs with innovative, high quality products and services at a competitive price. We are committed to sustained excellence in all dimensions of our business and to providing a fair return for our shareholders. Outstanding performance will provide the resources necessary for Northrop to grow as a world-class aerospace company.

OUR VALUES

We, the men and women of Northrop, are guided by the following values. They describe our company as we want it to be. We want every decision and action on the job to demonstrate these values. We believe that putting our values into practice creates long-term benefits for our shareholders.

We work to deliver CUSTOMER SATISFACTION ...

We each have valued customers, whether they be Government agencies, other businesses, our shareholders, or one another. We believe in respecting our customers, listening to their requests, understanding their expectations, and delivering products and services that meet agreed-upon standards.

We value NORTHROP PEOPLE ...

We treat one another with respect, recognizing the significant contributions that come from the diversity of individuals and ideas. Clear expectations must be communicated and people should know how they are doing. We create an environment of teamwork, openness, challenge, and development.

We regard our SUPPLIERS as essential team members...

We owe our suppliers the same type of respect that we show to our customers. Our suppliers deserve fair and equitable treatment, clear agreements, and honest feedback on performance. We consider our suppliers' needs in conducting all aspects of our business.

We take responsibility for the QUALITY of our work ...

We will deliver excellence and continue to strive for improvement. We define Quality as "meeting customer requirements—including cost and schedule." Each of us is accountable for the quality of whatever we do.

We demonstrate INTEGRITY in all we do ...

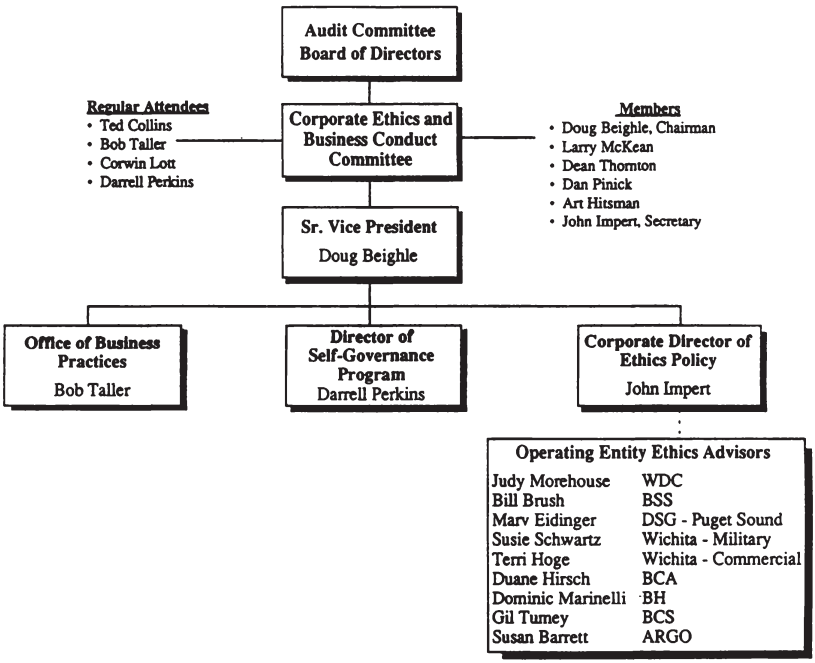
We are each personally accountable for high standards of behavior, including honesty and fairness in all aspects of our work. We strive to fulfill our commitments as law-abiding citizens and conscientious employees. We will be responsible and reliable, truthful and accurate, and diligent in the effective use of company and customer resources.

We provide LEADERSHIP as a company and as individuals ...

Northrop's leadership is founded on advanced technology, innovative manufacturing and sound business management. Our continued success requires developing people and rewarding their achievements. We each lead through competence, creativity, and teamwork.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Schaubild 2



SG-002 Figure 2

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Schaubild 3

How We Will Send the Messages — To Whom and When

Responsible Parties: Corporate Employee Communications/ Ethics & Business Conduct	**Media:**	**Test Market Themes and Designs**	**Kresa Letters**	**Company Annual Report**	**Ethics Report (Spring)**	**"How-To" Guide (proposed)**
	Audiences:	• Selected managers • Selected employees	• All managers • All employees	• Shareholders • Board of Directors • Managers • Employees • Community	• All managers • All employees • Compliance agencies	• All managers • All employees
Responsible Parties: Division and Corporate Departments	**Media:**	**"When Things Went Wrong" Video (crossfeed 1st and 2nd quarters)**	**1991 Recap**			**2nd Crossfeed Video (proposed)**
	Audiences:	• Senior managers • All managers	• Board of Directors meetings • Corporate Policy Council meetings • VP/GM staff meetings			• Senior managers • All managers
Responsible Parties: Division Employee Communications/ Ethics & Business Conduct	**Media:**				**Divisional GM/VP Letters**	
	Audiences:				• All Division employees	
Responsible Parties: All Managers	**Media:**				**Special Presentations (proposed)**	
	Audiences:				• All direct reports • All Div/Dept staff	
	Timeline	**February**	**March**		**Second Quarter**	**Third/Fourth Quarters**

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

"Northrop Knowledge" Game (proposed) • Inter-division tournament • All managers • All employees	**Leadership Conference Inventory** • Senior executives – Round 2 1/92 – Round 3 1/92 • Managers – Round 2 9/92-2/93	**Leadership Inventory Phone Calls (quarterly)** • Senior executives • Managers	**Bulletins (monthly – begin January '92)** • All managers	**Northrop News** • All employees	**On-Line Information Service (proposed 1993)** • All managers • All employees	
	Executive Visibility Program • Selected managers • Selected employees	**Values Challenge Forums with Executives (proposed)** • Selected managers • Selected employees	**Baldrige Assessment Leadership Plans** • Corporate Policy Council • VP/GMs • Managers	**Community/ Industry Forums** • Community • Compliance agencies • Suppliers	**Journal/Trade Publications** • Community • Suppliers • Compliance agencies • Employees	**Compliance Documents** • DCA • DCAA • DCMC
	Divisional Newsletters • All employees	**Personal Interaction** • Managers • Employees • Community • Compliance agencies • Suppliers	**Values Teams** • Selected managers • Selected employees	**Posters** • All employees	**Upward Feedback Mechanisms** • All employees	**Training and Development Programs** • Managers • Employees
	Face-to-Face Contact • Direct reports • Div/Dept employees	**Management Meetings** • All direct reports • All Div/Dept managers and employees	**Annual Business Plans** • Senior management • All Div/Dept managers and employees	**Performance Management Process (quarterly)** • All direct reports		
1993	Ongoing					

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Organisationales Lernen — Was kann die ökonomische Theorie dazu beitragen?

Von *Margit Osterloh, Simon Grand* und *Regine Tiemann,* Zürich

1. Zur Relevanz des organisationalen Lernens

Die Globalisierung der Märkte, der Siegeszug der Informationstechnologien und die Verkürzung der Produktlebenszyklen bei gleichzeitiger Verlängerung der Produktentwicklungszeiten stellen die Unternehmen vor gänzlich neue Herausforderungen. Nachhaltige Wettbewerbsvorteile werden nur dann erzielt, wenn neben die kostengünstige Produktion gleichzeitig die Fähigkeit zur schnellen Reaktion auf sich rasch ändernde Märkte tritt.

Als Antwort auf diese neuen Anforderungen ist eine Flut neuer Businesskonzepte wie Lean Management (*Womak / Jones / Roos* 1990), Prozeßorganisation (*Hammer / Champy* 1993; *Davenport* 1993) oder Virtual Corporation (*Davidow / Malone* 1992) entstanden. Das Gemeinsame dieser Konzepte ist die Betonung von Teamproduktion, Abflachung der Hierarchien und „Empowerment" (d.h. größere Entscheidungsspielräume vor Ort). Im Gegensatz zu herkömmlichen tayloristischen Organisationen wird die Trennung von Entscheiden und Ausführen aufgehoben. Statt dessen wird „knowledge work" überall im Unternehmen erforderlich.

In der wissenschaftlichen Diskussion findet diese Entwicklung ihren Niederschlag in der Debatte um organisationales Lernen und Kernkompetenzen. Diese Debatte hat ihren theoretischen Hintergrund in einem neuen Paradigma des strategischen Managements, dem resource-based-view bzw. dem dynamic-capability-view. Dieses neue Paradigma wird dem herkömmlichen market-based-Paradigma gegenübergestellt, welches in der Industrial Organization-Tradition steht:[1] Dem market-based-view wird vorgeworfen, daß er die Unternehmensressourcen als Ursache von Wettbewerbsvorteilen analysiert, die *Dauerhaftigkeit, Nicht-Substituierbarkeit und Nicht-Imitierbarkeit der Ressourcen*[2] aber zu wenig

---

[1] Diese Gegenüberstellung ist eine Fortsetzung der bereits bekannten Debatte um „technology push"- (*Rosenberg* 1976) versus „demand pull"-Ansatz (*Schmookler* 1966), allerdings erweitert um die Einbeziehung von sogenannten „tacit competencies", welche insbesondere die Bedeutung von organisatorischen Routinen und des organisationalen Lernens betonen, vgl. *Nelson* (1991).

[2] Vgl. *Barney* (1991). Zu einer Erwiderung aus der Sicht des market-based-view vgl. *Porter* (1991) sowie zu einem Versuch diese beiden Sichtweisen unter Einbeziehung

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

berücksichtige. Diese liegen vor allem in organisatorischen Faktoren begründet, insbesondere aber in der Fähigkeit zu organisationalem Lernen und im Verfügen über handlungsrelevantes „tacit knowledge" (*Polanyi* 1967).

Organisationen werden in zunehmendem Ausmaß zu Institutionen, in denen die *Produktion von Wissen* zum zentralen Wettbewerbsvorteil wird — ein Sachverhalt, den die neoklassische Ökonomik bis heute nur wenig beachtet hat. Sie hat sich statt dessen weitgehend auf den Aspekt des *Tausches von Informationen* beschränkt (vgl. *Siegenthaler* 1993). Diese Kritik an der ökonomischen Theorie steht im Gegensatz zu der in der letzten Zeit stattfindenden (Re-)Ökonomisierung der Betriebswirtschaft- und Managementlehre (vgl. *Ordelheide* u.a. [Hrsg.] 1991). Deshalb ist es besonders wichtig, sich mit den spezifischen Defiziten der Ökonomik in bezug auf die Erfassung des organisationalen Lernens und der Innovation auseinanderzusetzen, d.h. den Themen, die heute im Mittelpunkt betriebswirtschaftlichen Interesses stehen.

Dabei stehen zwei Aspekte im Vordergrund:

(1) Die Ökonomik unterscheidet nicht zwischen Information und Wissen *(Wissensaspekt)*:

Zwar beschäftigt sich die Ökonomik ausführlich mit Informationen und Informationsasymmetrien. Aber dabei werden Informationen wie Daten behandelt: Deren Erwerb reduziert Unsicherheit, und deren Mangel ermöglicht „shirking" oder opportunistisches Verhalten (*Alchian / Demsetz* 1972). Informationen werden dabei in der Ökonomik nicht anders als andere Güter angesehen, d.h., sie werden zu Gleichgewichtspreisen gekauft und verkauft. Wissen hingegen entsteht nicht allein durch Kauf von Informationen, sondern erst durch Verknüpfung dieser Informationen mit bereits vorhandenen kognitiven Konstrukten.[3] Dies bedeutet, daß Informationen erst dann zu Wissen transformiert werden, wenn sie auf dem Hintergrund von (Alltags-)Theorien interpretiert und als Bestandteil der persönlich verfügbaren Wissenbasis handlungsrelevant werden (vgl. *Nonaka* 1992).[4] Deshalb würde auch eine drastische Reduktion der Informationskosten (etwa durch den Einsatz moderner Informationstechnologien) keineswegs zu

---

spieltheoretischer Argumente zu vereinen (zu *Knyphausen* 1993). Allerdings geht zu Knyphausen so gut wie gar nicht auf den Aspekt des organisationalen Lernens ein.

[3] Kognitive Strukturen sind Wissensrepräsentationen, welche die neu gewonnenen Informationen organisieren. Diese Wissensrepräsentationen enthalten Kategorien, Konstruktionssysteme (z.B. Abstraktions- und Komplexionshierarchien, vgl. *Dörner* (1979), Kausalbeziehungen und Skripts (zeitliche Beziehungen zwischen Begebenheiten), vgl. *Schneider / Angelmar* (1993).

[4] Hierbei spielen neben den unterschiedlichen kognitiven Stukturen auch unterschiedliche kognitive Prozesse der Informationsverarbeitung (gekennzeichnet z.B. durch verschiedene Heuristiken) und verschiedene kognitive Stile (z.B. konvergentes und divergentes Denken) eine Rolle, vgl. *Schneider / Angelmar* (1993). Vgl. hierzu auch das Konzept des ipsativen Verhaltens von *Frey* (1990, S. 181 ff.) sowie die begriffliche Unterscheidung von semantischen und pragmatischen Informationen in der Semiotik.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

einem Verschwinden der Probleme des Wissenstransfers führen (vgl. *Minkler* 1993).

Weiter ist der Unterschied zwischen Information und Wissen[5] auch ursächlich dafür, daß in der neoklassischen Ökonomik nur der Fall der *parametrischen,* nicht der *strukturellen* Unsicherheit behandelt werden kann, obwohl der Unterschied durchaus geläufig ist:[6] *Parametrische* Unsicherheit entsteht, wenn der Entscheider ex ante alle möglichen Alternativen spezifizieren kann, aber nicht weiß, welche der Alternativen eintreten wird. *Strukturelle* Unsicherheit ist dann gegeben, wenn ex ante nicht bekannt ist, welche möglichen Alternativen existieren (vgl. *Minkler* 1993). Die weitgehende Beschränkung der neoklassischen Theorie auf parametrische Unsicherheit, d.h. auf die rationale Auswahl zwischen ex ante *gegebenen* Alternativen führt dazu, daß „the equilibrium concept is associated with a world view that treats the future as implied in the present" (*Buchanan / Vanberg* 1991, S. 170). Um ex ante nicht gegebene Alternativen überhaupt entwickeln zu können, brauchen die Individuen Theorien. Wenn aber verschiedene Individuen über verschiedene (Alltags-)Theorien verfügen, weil eben ein großer Teil des Wissens „tacit" ist, dann haben zwei Individuen selbst bei gleicher Information nicht dasselbe Wissen. Selbst vollständige Informationssymmetrie zwischen Prinzipal und Agent kann deshalb dazu führen, daß der Agent dem Prinzipal weiterhin überlegen ist, wenn er die bessere Theorie hat.[7]

(2) Die Ökonomik hat ein eingeschränktes Verständnis von der Vielfalt der institutionellen Mechanismen, die in Organisationen eingesetzt werden *(Institutionalisierungsaspekt)*:

Im Modell der neoklassischen Ökonomik entstehen Organisationen, weil die Transaktionskosten geringer sind als auf Märkten.[8] Die Kosten sind deshalb geringer, weil in Organisationen zentrale Entscheidungen gefällt werden und diese Entscheidungen innerhalb der arbeitsvertraglich festgelegten Indifferenzzone mittels Befehl durchgesetzt werden können (vgl. *Simon* 1951). Jedoch verfügen Organisationen nicht nur über ein reichhaltiges Inventar an Institutionalisierungs-

---

5 Genauer zwischen Daten und Wissen. Um Verwechslungen zu vermeiden, werden wir im folgenden den Begriff der Information immer im Sinne von Daten gebrauchen.

6 Vgl. die altbekannte Unterscheidung von Risiko und Unsicherheit bei *Knight* (1921).

7 Die Ökonomik diskutiert dieses Problem neuerdings unter dem Aspekt der „Entscheidungsanomalien". Vgl. z.B. *Frey* (1990), *Kirchgässner* (1991) und die Abschnitte 2.2.3 sowie 4.1 dieser Arbeit. Allerdings bemüht sie sich zumeist mit einer Reihe von „Immunisierungsstrategien", die Relevanz von Anomalien auf der Ebene von aggregierten Handlungen abzumildern, weil sie auf effizienten Märkten verschwinden würden, vgl. *Kirchgässner* (1991). Dies ist allerdings für die betriebswirtschaftliche Strategielehre keine befriedigende Perspektive, weil es ihre Aufgabe ist, ex ante dafür zu sorgen, daß Betriebe nicht aus dem Markt ausscheiden.

8 Die Ökonomik vernachlässigt damit den Orientierungsaspekt von Organisationen, d.h. den Sachverhalt, daß in Situationen struktureller Unsicherheit Organisationen Kommunikationsarenen darstellen, in denen Individuen sich wechselseitig zu subjektiv handlungsleitenen Alltagstheorien verhelfen, vgl. *Siegenthaler* (1993) unter Bezugnahme auf *Habermas* (1981).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

mechanismen zwischen den beiden Extremen Vertrag und Befehl,[9] welche neuerdings in der evolutionsökonomisch inspirierten Ökonomie unter dem Begriff „organisatorische Routinen" (*Nelson / Winter* 1982) Aufmerksamkeit finden.[10] Neben diesen organisatorischen Routinen, die den *Koordinationsaspekt* in Organisationen berücksichtigen, muß bei der organisatorischen Gestaltung auch noch dem *Motivationsaspekt* Rechnung getragen und mit dem Koordinationsaspekt in Übereinstimmung gebracht werden (vgl. *Frese* 1993). In bezug auf die Motivation hat die Ökonomik ein sehr eingeschränktes Verständnis von den Institutionalisierungsmechanismen in Organisationen. Sie stellt ausschließlich auf extrinsische Motivation ab, obwohl die Existenz von intrinsischer Motivation keineswegs geleugnet wird (vgl. z.B. *Milgrom / Roberts* 1991).[11] Insofern ist gerade die moderne Institutionenökonomik in bezug auf diesen wichtigen Aspekt der Institutionalisierungsproblematik unvollständig. Andererseits kann die Managementlehre mit Gewinn auf das reiche Inventar der Institutionenökonomik zurückgreifen, da sie ihrerseits bei der institutionellen Gestaltung von Anreizstrukturen zahlreiche Defizite aufweist.

Im Gegensatz zur Ökonomik hat die verhaltenswissenschaftliche Managementliteratur zu organisationalem Lernen den hier mit *Wissensaspekt* bezeicheten Unterschied von Information und Wissen ausführlich behandelt, wenngleich sie von einer geschlossenen Theorie des Wissenstransfers in Organisationen noch weit entfernt ist (vgl. *Huber* 1991). In bezug auf den *Institutionalisierungsaspekt* weist sie allerdings gegenüber der Institutionenökonomik einen starken Nachholbedarf auf. Der Aggregationsprozeß vom individuellen zum organisationalen Lernen wird zumeist nur auf der Ad-hoc-Ebene erörtert.[12] Zwar existieren Beschreibungen und Analysen darüber, wie organisatorische Strukturen das organisationale Lernen verbessern können (vgl. z.B. *Cohen* 1981; *Nonaka* 1988, 1992; *Starbuck* 1984; *Levitt / March* 1988). Aber die Befunde sind widersprüchlich (vgl. *Huber* 1991), und es fehlt fast gänzlich ein theoretisches Konzept, wie individuelles und organisationales Lernen systematisch aufeinander bezogen werden können.[13]

---

9 Diese bilden den Gegenstand der betriebswirtschaftlichen Organisationslehre. Solche Mechanismen sind z.B. die funktionale, die divisionale Organisation, die Matrix-, Projekt-, Guppen- oder Prozeßorganisation sowie zahlreiche Kombinationen dieser Grundformen.

10 Ohne daß dabei auf die existierende Organisationsliteratur Bezug genommen wird, vielmehr beschränkt man sich zumeist auf den Hinweis, daß organisatorische Routinen eine wichtige Quelle der „dynamic capabilities" sind, vgl. *Nelson* (1991).

11 Auch hier werden Immunisierungsstrategien angewendet, um intrinsische Motivation als „weak links" (*Mayer* 1993) behandeln zu können: Man müsse bei der Organisationsgestaltung auf den „worst case" abstellen (vgl. *Milgrom / Roberts* 1991, S. 42 ff.). Dabei wird außer acht gelassen, daß man damit sich selbst erfüllende Prophezeihungen erzeugt, indem man die bestehende intrinsische Motivation zerstört, vgl. zu diesem Aspekt Abschnitt 4.4 dieser Arbeit.

12 Zu den wenigen Ausnahmen gehören die Vertreter des soziologisch inspirierten Institutionenansatzes, vgl. z.B *Meyer / Rowan* (1977).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Neuerdings betont der von der Evolutionsökonomie beeinflußte resource-based-view bzw. dynamic-capability-view die Notwendigkeit, den Wissensaspekt mit dem Aspekt der Institutionalisierung von Lernen zu verbinden. Die Autoren tun dies, indem sie Organisationen als eine Hierarchie von organisatorischen Routinen kennzeichnen, welche zu dauerhaften „core capabilities" führen (vgl. *Nelson / Winter* 1982; *Nelson* 1991; *Teece / Pisano / Shuen* 1990; *Minkler* 1993). Wenn wir im folgenden auf diese Literatur nicht weiter eingehen werden, dann deshalb, weil dort zwar immer wieder der Wissenstransfer durch organisatorische Routinen als zentral bezeichnet wird. Es wird aber nur wenig zur Explikation beigetragen, wie man sich den Institutionalisierungsprozeß für den Wissenstransfer vorzustellen hat.

Wir wollen deshalb im folgenden in erster Linie die verhaltenswissenschaftliche Managementliteratur und die neoklassische Ökonomik gegenüberstellen, um herauszufinden, inwieweit beide Theorien in bezug auf den *Wissensaspekt* und den *Institutionalisierungsapekt* voneinander lernen können. Wir beschränken uns dabei auf die theoretischen Aspekte dieser Integration. Dies kann für die Managementwissenschaft als einer anwendungsorientierten Wissenschaft allerdings nur der erste Schritt sein, dem Gestaltungsempfehlungen folgen müssen.

2. Organisationales Lernen aus der Sicht der psychologisch orientierten Managementlehre

2.1 Vorbemerkung

Unter organisationalem Lernen versteht man eine unternehmensspezifische Form des Wissenserwerbs, welche unabhängig vom Wechsel der Mitarbeiter ist. Die meisten Konzepte des organisationalen Lernens gehen vom individuellen Lernen aus, ohne zu klären, wie die einzelnen individuellen Lernprozesse zu einem organisationalen Lernprozeß aggregiert werden können: Organisationen werden als eigene Entität reifiziert. Bücher wie „The Thinking Organization" (*Sims / Gioia* 1986) oder „Die lernende Organisation" (*Sattelberger* [Hrsg.] 1991) benutzen individuelles Lernen als Analogie oder Metapher für organisationales Lernen. Darüber hinaus klären sie auch inhaltlich nicht, „what is organizational about cognition and what is cognitive about organizations" (*Schneider / Angelmar* 1993, S. 348). Die daraus resultierende Reifizierung ist wenig hilfreich. Um sie zu vermeiden, wollen wir zunächst einen kurzen Überblick über die in der

[13] Dies führt zu fragwürdigen Meßinstrumenten: So wird organisationales Lernen entweder mittels Durchschnittsbildung aus individuellem Lernen abgeleitet, oder es werden globale Maße, z.B. Kenntnisse eines Organisationsmitgliedes in leitender Stellung herangezogen. In beiden Fällen wird das eigentlich Organisatorische am Lernen, nämlich wie Wissen kommuniziert wird, im Meßverfahren unberücksichtigt gelassen, vgl. *Schneider / Angelmar* (1993).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Managementlehre gängigen Konzepte des individuellen Lernens geben und erst dann darauf eingehen, was das organisationale vom individuellen Lernen unterscheidet.

2.2 Konzepte des individuellen Lernens

Die Managementlehre greift in ihren Konzepten auf psychologische Lerntheorien zurück. In der Psychologie werden zwei Gruppen von Theorien des individuellen Lernens unterschieden: behavioristische und kognitive Lerntheorien. Der Behaviorismus stellt Lernen als Stimulus-Response-Modell dar, d.h., Verhalten wird vollständig durch eine Abfolge von Reiz-Reaktionsmechanismen bestimmt. Der Behaviorismus stellt deshalb den Übergang von Information zu Wissen nicht dar.

Um dieses Defizit auszugleichen, erweitern Theorien des kognitiven Lernens die behavioristischen Lerntheorien um den Prozeß der Wissensgenerierung.[14] Dadurch wird die Entwicklung und Veränderung von kognitiven Strukturen thematisiert. Kognitive Strukturen sind nach *Piaget* (1973) generalisierte Formen der internen Organisation des Handelns, welche im Sinne verallgemeinerter Regeln die Handlung steuern. Vergleichbare Konstrukte finden sich hinter den Begriffen „Schema" (*Bartlett* 1932), „cognitive maps" (*Tolman* 1948), „frame" (*Minsky* 1968; *Tversky / Kahneman* 1981), „script" (*Schank / Abelson* 1977) oder „building blocks of cognition" (*Rumelhart* 1980).[15] Es geht dabei um die kontextspezifische Bedeutungsstiftung (Informationsverarbeitung) und Problemlösung (vgl. *Frey,* D. 1987, S. 51 ff.). Die Gesamtheit der Schemata, „cognitive maps", „frames" etc., über die ein Mensch verfügt, stellt sein Wissen im Sinne seines Handlungsrepertoires oder seiner Handlungskompetenz dar.[16]

Soweit sich die Theorien des organisationalen Lernens auf individuelles Lernen beziehen, stützen sie sich explizit oder implizit auf kognitive Lerntheorien.[17] Dies wird insbesondere an der Auseinandersetzung mit den Begriffen „frames", „cognitive maps" etc. deutlich (vgl. z.B. *Stata* 1989, S. 64; *Senge* 1990, S. 174 ff.;

---

[14] Die Integration behavioristischer in kognitivistische Lerntheorien wird besonders deutlich im TOTE (test-operate-test-exist)-Modell, das *Miller / Galanter / Pribram* (1960) in Analogie zum kybernetischen Regelkreismodell entwickelten. Das TOTE-Schema ersetzt das behavioristische Reflex-Modell.

[15] Als Überblick vgl. *Schwarz* (1985).

[16] An diese Tradition knüpfen auch neuere Entwicklungen in der Cognitive Science an; vgl. *Winograd / Flores* (1992), *Simon* (1991), *Maturana / Varela* (1984), *Suchman* (1987), und neuerdings auch *Denzau / North* (1994).

[17] So stellt beispielsweise *Weick* (1991) explizit behavioristische und kognitive Lerntheorien gegenüber und zeigt, daß die kognitivistische Perspektive für das Phänomen des organisationalen Lernens angemessener ist. Er kritisiert allerdings an den kognitivistischen Theorien, daß sie eine quasi automatische Umsetzung von Kognition in Handlungen unterstellen (vgl. ebd. S. 122). Diese Kritik gilt nicht für das Modell von *Neisser* (1979), das wir in Abschnitt 2.2.2 darstellen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Hedberg 1981, S. 3; *deGeus* 1988, S. 74; *Levitt / March* 1988, S. 324; *Schneider / Angelmar* 1993, S. 349). Im folgenden wollen wir einen kurzen Überblick über Konzepte geben, welche die Veränderung von kognitiven Strukturen thematisieren und als Grundlage des organisationalen Lernens in der Managementliteratur diskutiert werden.

2.2.1 Der Assimilations-Akkommodations-Zyklus nach Piaget

Piaget beschreibt die Veränderung von kognitiven Strukturen im Lernprozeß von Kindern und Jugendlichen als Wechselwirkung von Assimilation und Akkomodation. Assimilation bezeichnet den Sachverhalt, daß Individuen die Umwelt gemäß ihren vorhandenen kognitven Schemata begreifen. Akkomodation meint die Veränderung der bestehenden Strukturen, wenn diese einen erfolgreichen Umgang mit der Umwelt nicht mehr gewährleisten. Piaget sieht dabei allerdings die Umwelt vorwiegend unter dem Blickwinkel mathematisch-physikalischer Erkenntnis: Die Welt ist dem Individuum vorgegeben und kann nicht handelnd verändert werden. Daher ist sein Assimilations-Akkomodations-Modell von einem gleichsam kybernetischen Mechanismus gekennzeichnet (zu dieser Kritik vgl. *Keller* 1976).

2.2.2 Der Erkenntniszirkel nach Neisser

Neisser entwickelt die kognitive Lerntheorie von Piaget weiter, indem er Aspekte des sozialen Handelns einbezieht. Er geht davon aus, daß das Individuum seine Umwelt entsprechend seinen kognitiven Strukturen handelnd verändert, so daß der Assimilations-Akkommodations-Zyklus um die Stufe der Exploration erweitert wird (vgl. Abb. 1).

Die Assimilation enspricht dem Selektionsprozeß, die Akkomodation dem Modifikationsprozeß. Zusätzlich zu Assimilation und Akkomodation wird bei Neisser der Vorgang der Exploration eingeführt, ein Vorgang, in dem die Situation handelnd verändert wird. Dieser kann durch alle am sozialen Prozeß beteiligten Individuen initiiert werden. So kann Neisser den Vorgang der Akkomodation als aktiven, sozialen Prozeß modellieren, während dieser bei Piaget naturgesetzlich vorgegeben ist. Folgt man dieser Überlegung, dann wäre damit ein erster Hinweis auf eine wichtige Schnittstelle zwischen sozialem Lernen und individuellem Lernen gegeben, ohne die Organisation als Entität zu reifizieren.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

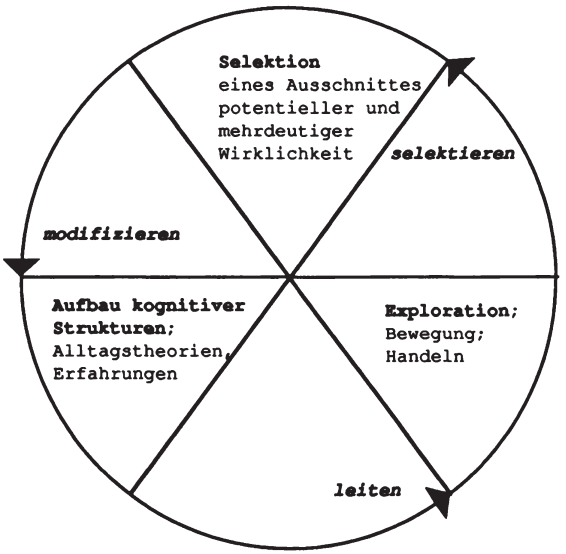


Abb. 1: Erkenntniszirkel nach Neisser

2.2.3 Die Prospekttheorie nach Tversky / Kahneman

Tversky / Kahneman (1981, 1984) zeigen in ihrer Prospekttheorie experimentell auf, daß es bei der Beeinflussung der von Piaget und Neisser behandelten „frames" (bzw. Schemata) empirische Regelmäßigkeiten gibt, die nicht nur im Lernprozeß des Kindes wirksam sind. Sie können deshalb zur Beeinflussung erwachsener Personen in sozialen Kontexten systematisch genutzt werden.

Prospekte bezeichnen Alternativen, bei denen auch unsichere Resultate (Risiken) vorkommen können. In den Experimenten konnte gezeigt werden, daß die Einschätzung der Risiken systematisch beeinflußt werden kann, beispielsweise indem bestimmte Problemformulierungen gewählt werden:[18] So wird eine Wirtschaftspolitik, welche die Arbeitslosenquote von 10 auf 5% reduziert, besser eingeschätzt als eine, welche die Beschäftigungsquote von 90 auf 95% erhöht.

Diese Ergebnisse bedeuten Abweichungen von der Hypothese der Erwartungsnutzenmaximierung der Ökonomik und werden deshalb auch als Entscheidungsanomalien bezeichnet. Im Zusammenhang des organisationalen Lernens sind sie deshalb von Bedeutung, weil sie einen Hinweis darauf geben, wie durch die systematische Präsentation von Informationen Entscheidungsprozesse beeinflußt werden.[19]

---

[18] Ähnliche Effekte werden seit langem als das sogenannte Allais-Paradoxon diskutiert, vgl. als Übersicht *Frey / Eichenberger* (1989) und *Kirchgässner* (1991, S. 144 ff.).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

2.2.4 Lernniveaus nach Bateson

Bei Tversky / Kahneman wird die Herkunft der „frames" als ein idiosynkratischer Prozeß beschrieben (vgl. dazu die Kritik von *Lindenberg* 1990), die systematische Reflexion über die Wahl angemessener „frames" ist nicht Gegenstand der Prospekttheorie und kann im Rahmen einer experimentellen Psychologie auch kaum untersucht werden. Dazu wäre die Einführung verschiedener Reflexionsniveaus nötig.

Ein prominenter Ansatz dazu wurde von *Bateson* (1992) geliefert. Er unterscheidet die Lernniveaus im Hinblick auf eine Klassifikation von Irrtumstypen, die in verschiedenen Lernprozessen korrigiert werden können. *Lernen 0* ist auf behavioristisches Lernen beschränkt. *Lernen I* umfaßt die Auswahl zwischen verschiedenen Alternativen, die in einem Kontext vorgegeben sind. Es setzt die Einbettung der Alternativen in einen konzeptuellen Rahmen bzw. in eine (Alltags-)Theorie voraus. Bei *Lernen II,* auch Deutero-Lernen[20] genannt, geht es um die korrigierende Veränderung der Menge von Alternativen, unter denen ein Individuum seine Auswahl trifft. Hier geht es um den Vergleich und die kritische Abwägung zwischen verschiedenen Konzepten oder Theorien. Durch *Lernen III* wird der Identitätsbildungsprozeß des Individuums verändert.[21] Die Lernniveaus nach Bateson bringen damit zum Ausdruck, daß Lernen oder der Erwerb von Wissen erst durch die Interpretation von Informationen im Lichte von Theorien zustandekommt und daß diese Theorien auf verschiedenen Reflexionsniveaus angesiedelt sein können. Der Fortschritt des Lernens ist dabei als ein irreversibles Fortschreiten von Stufe zu Stufe gekennzeichnet: Ist ein Individuum auf einem höheren Lernniveau angelangt, dann kann es zwar erklären, warum es auf dem niedrigeren Lernniveau anders geurteilt hat, es kann sich aber nicht auf eine niedrigere Stufe zurückversetzen.

Die empirische Basis des Konzeptes von Bateson ist allerdings umstritten. Es handelt sich um Ergebnisse von Einzelfallstudien auf psychoanalytischem Hintergrund. Gleichwohl bildet es die Grundlage für die meisten Konzepte des organisationalen Lernens.

---

19 Diese Erkenntnisse finden mittlerweile auch Eingang in die betriebswirtschaftliche Theorie der Strategiebildung, vgl. *Bateman / Zeithaml* (1989).

20 Deutero = griech. der zweite.

21 Lernen III wird im Rahmen des organisationalen Lernens nicht behandelt. Vgl. im übrigen auch die in der Problemlösungstheorie übliche Klassifikation von Entscheidungssituationen in Algorithmen, geschlossene Probleme, offene Probleme und dialektische Probleme, z. B. bei *Dörner* (1979), *Osterloh* (1991). Diese Klassifikation umfaßt feinere Abstufungen zwischen Lernen 0 und Lernen II.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

2.3 Organisationales Lernen

Die meisten Konzepte zum organisationalen Lernen schließen an die Idee der verschiedenen Lernniveaus an. Allerdings ist die Abgrenzung der Lernniveaus nicht einheitlich (vgl. z.B. *Fiol / Lyles* 1985; *Shrivastava* 1983). Am bekanntesten ist die Konzeption von *Argyris / Schön* (1978) geworden. Sie unterscheiden in Anlehnung an *Bateson* (1992) zwischen Single-loop-, Double-loop- und Deutero-Lernen. *Single-loop-Lernen* beschreibt das Lernen innerhalb gegebener „frames", *Double-loop-Lernen* das Verändern und die Entwicklung von neuen „frames", *Deutero-Lernen* den reflektierten Umgang mit Single-loop- und Double-loop-Lernen.

Die Verbindung zwischen individuellem und organisationalem Lernen wird bei Argyris / Schön durch die „theories of action" hergestellt.[22] Diese umfassen „norms, strategies, and assumptions. It includes in its scope the organization's patterns of communication and control, its ways of allocating resources to goals, and its provisions for self-maintenance" (*Argyris / Schön* 1978, S. 15). In der Literatur zum organisationalen Lernen findet man anstelle des Begriffs „theories of action" auch die Begriffe „kognitive Strukturen", „frames", „mental maps" oder „cognitive models" (vgl. oben), die zum Teil um den Zusatz „shared" ergänzt werden. Dies weist nochmals darauf hin, daß organisationales Lernen immer wieder in Metaphern des individuellen Lernens dargestellt wird (vgl. *Dodgson* 1993, S. 380), ohne daß der Aggregationsprozeß vom individuellen zum organisationalen Lernen erklärt wird.

2.3.1 Single-loop-Lernen

Das Single-loop-Lernen kann man sich als einen einfachen Feedback-Prozeß vorstellen, in dem die gemeinsam geteilten „frames" ständig angewendet und dabei reproduziert werden (vgl. auch *Giddens* 1988). Informationen aus der Unternehmensumwelt werden auf der Basis der „shared mental maps" interpretiert und in routinisierte Verarbeitungsmechanismen überführt. Die Bedeutung der Routinen besteht darin, Stabilität beim Erwerb und bei der Interpretation von Wissen zu erzeugen: „the organizations's ability to remain stable in a changing context denotes a kind of learning" (*Argyris / Schön* 1978, S. 18).

Für einige Autoren wie *Levitt / March* (1988, S. 320)[23] beruht das Lernen in Organisationen sogar ganz wesentlich auf Routinen: „organizations are seen as learning by encoding inferences from history into routines that guide behavior.

---

[22] Die „theories of action" sind nicht als soziologische Handlungstheorien zu verstehen, sondern als organisationale „frames". Sie entsprechen nach *Pautzke* (1989, S. 105) dem Begriff der Unternehmenskultur (vgl. z.B. auch *Schein* 1986).

[23] In Anlehnung an *Cyert / March* (1963) und *Nelson / Winter* (1982).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

The generic term ‚routines' includes the forms, rules, procedures, conventions, strategies, and technologies around which organizations are constructed and through which they operate" (ebd.). Organisatorische Routinen haben die Eigenschaft, vom Wechsel einzelner Individuen unabhängig zu sein (vgl. ebd.; *Schneider / Angelmar* 1993, S. 355 in Anlehnung an *Weick* 1969; *Dodgson* 1993, S. 384). Offen bleibt in dieser Argumentation, wie Routinen auf dem Hintergrund gemeinsam geteilter „frames" gelernt und reproduziert werden.

2.3.2 Double-loop-Lernen

Nach *Argyris / Schön* (1978) ist die Veränderung des „frames" Gegenstand des organisationalen Double-loop-Lernens, das man sich als einen doppelten Feedback-Prozeß vorstellen muß. In diesem Prozeß werden neue Sachverhalte mit dem bestehenden „frame" verglichen. Bei nicht überbrückbaren Diskrepanzen zwischen beiden wird eine Veränderung des „frames" eingeleitet. Das Organisationale an diesem Vorgang wird dadurch akzentuiert, daß es sich hierbei um einen interindividuellen Konfliktlösungsprozeß handelt (vgl. *Argyris / Schön* 1978, S. 20 ff.): Die Organisationsmitglieder müssen sich über verschiedene Problemdefinitionen verständigen und eine gemeinsame Situationsdefinition generieren (vgl. *Senge* 1990a). In der Literatur zum organisationalen Lernen wird üblicherweise angenommen, daß dies nur über Kommunikation geleistet werden kann (vgl. z.B. *Senge* 1993; *Isaacs* 1993; *Schein* 1993).[24]

Ausführlich wird die Frage diskutiert, warum Double-loop-Lernen immer wieder verhindert wird. Der Prozeß der Neudefinition von „frames" ist in der Regel mit tiefgreifenden Verunsicherungen verbunden, die sich zu organisationalen Krisen ausweiten können (vgl. *Nystrom / Starbuck* 1984; *Fiol / Lyles* 1985, S. 808).[25] Diese krisenhaften Erscheinungen werden in anderem Zusammenhang auch unter dem Begriff „groupthink" (*Janis* 1982) behandelt: Es gibt Situationen, in denen das Gruppendenken individuelles Denken überlagert, unterdrückt oder sogar gänzlich verhindert (vgl. z.B. *McCaskey* 1991, S. 139, 141; *Schneider / Angelmar* 1993, S. 354). Gruppendenken kann denn auch als wechselseitiges „framing" im Sinne der Prospekttheorie interpretiert werden.[26]

---

[24] Diese Aussage wird im Zusammenhang mit spieltheoretischen Überlegungen relativiert werden müssen.

[25] Dies wird neuerdings auch in der ökonomischen Literatur akzentuiert, etwa wenn *Siegenthaler* (1993) von Prozessen des fundamentalen Lernens spricht.

[26] Diese — nicht nur negativen — Prozesse des „groupthink" werden in der ökonomischen Gruppentheorie (*Olson* 1968; *Sandler* 1992) gänzlich vernachlässigt. Die Einbeziehung von Gruppendenken und Gruppenkohäsion ist die Hauptdifferenz zwischen ökonomischen und sozialpsychologischen Gruppenkonzepten.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

2.3.3 Deutero-Lernen

Im Deutero-Lernen geht es um den reflektierten Umgang mit Single-loop- und Double-loop-Lernprozessen, z.B. um die Entscheidung, wann Singe-loop- und wann Double-loop-Lernen eingeleitet werden soll (vgl. *Argyris / Schön* 1978, S. 42, 308 ff.). Wie genau Deutero-Lernen im organisatorischen Kontext umgesetzt werden soll, wird in der Literatur nur vage angedeutet. Es geht beispielsweise um das Schaffen offener Argumentationskulturen und kreativer Freiräume. Auch wenn das Konzept des Deutero-Lernens äußerst unscharf definiert ist, so kann man sich doch anhand eines berühmt gewordenen Beispiels verdeutlichen, was für Konsequenzen es haben kann, wenn Deutero-Lernen nicht stattfindet:

Die Challenger-Katastrophe im Januar 1986 gehört zu den am besten dokumentierten Beispielen (vgl. *Presidential Commission* 1986) der negativen Auswirkungen von „groupthink" und zugleich der Verhinderung von Deutero-Lernen.[27] Der Space Shuttle war wenige Sekunden nach dem Start explodiert, weil wichtige Dichtungsringe aufgrund der tiefen Außentemperaturen spröde geworden waren. Obwohl dieses Problem sowohl der Nasa als auch der Herstellerfirma der Dichtungsringe bekannt war, wurde der Start freigegeben. Die Analyse des Verlaufs der Entscheidung zeigt, daß die Warnungen von zwei Ingenieuren vom Management übergangen wurden. Darüber hinaus wurde sogar aus dem Katalog für die wichtigsten Abbruchkriterien das Minimumkriterium für die Temperaturen gestrichen. Bei der entscheidenden Videokonferenz 15 Stunden vor dem Start fiel der Satz: „Take off your engineering hat and put on your management hat" (zitiert nach *Löhr* 1991, S. 17). Die Manager waren offensichtlich nicht in der Lage, die beiden handlungsleitenden „frames" des „managers" und des „engineers" auf einer Deutero-Ebene miteinander zu verbinden.[28]

2.4 Vorläufiges Fazit

Wie wir gesehen haben, werden in Theorien zum organisationalen Lernen die Individuen als Träger von Lernprozessen angesehen. Der Aggregationsprozeß vom individuellen zum organisatorischen Lernen wird selten expliziert. Die Folge davon ist, daß Organisationen entweder holistisch als „Lernen der Organisation" interpretiert werden, beispielsweise wenn behauptet wird: „Organizations learn from experience" (*March / Sproull / Tamuz* 1991, S. 1); oder wenn individuelles Lernen als Metapher für das organisationale Lernen herangezogen wird. Beides sind unbefriedigende Alternativen.

---

[27] Zur Interpretation der Challenger-Katastrophe unter dem „groupthink"-Aspekt vgl. *Moorhead / Ference / Neck* (1991).

[28] Auf die ethische Relevanz dieses Rollenkonfliktes zwischen „manager" und „engineer" geht *Löhr* (1991) ein.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Im folgenden wollen wir versuchen, mit Hilfe der ökonomischen Theorie organisationales Lernen auf der Basis eines methodologischen Individualismus abzuleiten.[29] Gleichzeitig wollen wir untersuchen, wo die Grenzen des ökonomischen Ansatzes bei der Bewältigung der eingangs erwähnten praktischen Problemstellungen liegen.

3. Institutionalisierung als „missing link" zwischen individuellem und organisationalem Lernen

3.1 Zum Begriff der Institutionalisierung

Will man Organisation nicht als eigene, von den Individuen losgelöste Entität reifizieren, so darf man nicht von einer „Lernenden Organisation" (vgl. z.B. *Sattelberger* [Hrsg.] 1991) sprechen. Eine Organisation kann nicht lernen. Dies können nur Individuen (vgl. *Simon* 1991, S.126). In einer Organisation als Set von dauerhaften Regelungen kann allerdings dafür gesorgt werden, daß Prozesse des individuellen Lernens zielgerichtet institutionalisiert werden.

Der Institutionenbegriff ist nun allerdings selbst vieldeutig, auch innerhalb einer am methodologischen Individualismus orientierten Sozialwissenschaft.[30] Zwar kann ein allgemeines Einverständis darüber vorausgesetzt werden, daß Institutionen den Zweck haben, die Unsicherheit menschlicher Interaktion zu vermindern (*North* 1992, S. 30), indem sie „Spielregeln einer Gesellschaft" (ebd. S. 3), „soziokulturelle Rahmenbedingungen individuellen Handelns" (*Vanberg* 1983, S. 55) oder „aufeinanderbezogene Verhaltensnormen" (*Gäfgen* 1983, S. 20) schaffen. In diesem Sinne sind Institutionen nicht bloß organisierte soziale Gebilde, sondern Komplexe von Regelungen für die Interaktion zwischen Individuen, wie z.B. Geld, Eigentum, Recht, Markt, Preissystem oder Traditionen (vgl. *Vanberg* 1983, S. 55 ff.; *Frey* 1990, S. 2 ff.). Ebenso dürfte es unkontrovers sein, von Institutionalisierung als der Erzeugung von systematisch reproduzierten Handlungsmustern zu reden (vgl. *Giddens* 1988). Aber in bezug auf die Frage, warum solche Handlungsmuster systematisch reproduziert werden, existieren

---

[29] Methodologischer Individualismus wird hier als Gegensatz zum methodologischen Holismus verstanden. Der methodologische Individualismus umfaßt zwei Varianten, die dezisionistische und die nichtdezisionistische. In der dezisionistischen Variante, welche innerhalb der Ökonomik vorherrscht, sind Präferenzen gegeben und werden nicht weiter erklärt. In der nichtdezisionistischen Variante, zu welcher der Symbolische Interaktionismus sowie die Habermas'sche Diskurstheorie zu rechnen sind, werden Präferenzen nicht als gegeben vorausgesetzt, vielmehr entstehen diese durch Kommunikation und interpretative Bedeutungszuweisungen (vgl. *Osterloh* 1994).

[30] Zu den Ebenen wissenschaftlicher Theoriebildung und deren Verknüpfung mit betriebswirtschaftlicher Praxisberatung vgl. *Osterloh / Grand* (1994). Zu einem praxisorientierten Ansatz der Umsetzung der Idee der organisatorischen Routinen vgl. das Konzept des „business process redesign" bei *Davenport* (1993). Als Übersicht über dieses und verwandte Konzepte vgl. *Osterloh / Frost* (1994).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

gegensätzliche Sichtweisen. Auf diese wollen wir uns im folgenden konzentrieren. Wir behandeln dabei nicht spezifische Ausprägungen von konkreten institutionellen Regeln in Organisationen, wie dies Gegenstand der betriebswirtschaftlichen Organisationslehre ist, sondern wir fragen grundsätzlich, warum Individuen Spielregeln akzeptieren bzw. warum Normen Geltung erhalten.

3.2 Normgeltung als zentraler Begriff

Einem Vorschlag von *Kliemt / Zimmerling* (1993) zufolge kann man bezüglich der Normgeltung zwischen Geltung 1 und Geltung 2 unterscheiden. *Geltung 1* besitzt eine Norm, wenn sie in einer Gesellschaft regelmäßig aufgrund von Sanktionen befolgt wird. Dabei kommt es nur auf die regelmäßige Verknüpfung von Sanktion und Normeneinhaltung an, nicht darauf, ob sich die Individuen bei der Normbefolgung auch den Norminhalt vorstellen. Sie handeln *normenkonform,* aber nicht *aufgrund* der Norm.

Bei *Geltung 2* hingegen gehört die Norm zum Vorstellungsgehalt der Individuen. Sie wird zum Handlungsgrund ähnlich einer „Verfassung", die sich Individuen geben. Geltung 2 unterscheidet sich von Geltung 1 wie die „Wahl einer Verfassung" von „Handeln unter einer Verfassung" oder wie die „Wahl *von* Regeln" von der „Wahl *nach* Regeln" (vgl. *Vanberg* 1989). Allerdings bedeutet die Wahl einer Verfassung oder Regel nicht, daß ausnahmslos nach dieser Regel gehandelt wird. Insofern schließt Geltung 2 den Gebrauch von Sanktionen im Sinne von Geltung 1 nicht aus.[31] Entscheidend ist, daß Geltung 2 in einigen Fällen zu „nichtstrategischem Handeln" oder „Regelbefolgung ohne Kalkulation" führt.

Geltung 2 kann zwei Ausprägungen annehmen: die Wahl von Faustregeln, Konventionen oder Routinen (Geltung 2a) und die Wahl von verpflichtenden Regeln (Geltung 2b). Die Wahl von Faustregeln oder Routinen *(Geltung 2a)* hängt eng mit dem „satisficing behavior" (*Simon* 1961) zusammen. Faustregeln geben angesichts komplexer Verhältnisse Verhaltenshinweise aufgrund von vereinfachten Situationsdefinitionen, welche als ganzes Set aktiviert werden. Die Verhaltensmuster können als vereinfachte deontologische Regeln auftreten, beispielsweise wenn der Kategorische Imperativ von Kant auf die jedem Kind vertraute Regel reduziert wird: „Was du nicht willst, was man dir tu, das füg auch keinem anderen zu". Sie können aber auch aufgrund von Kalkulationen entstanden sein, z.B. bei der Wahl zwischen verschiedenen Rollenmustern. Entscheidend ist, daß *nach* der Wahl des Verhaltensmusters in der konkreten Entscheidungssituation nicht mehr bewußt reflektiert resp. kalkuliert wird (*Kliemt / Zimmerling* 1993, S. 13).

---

[31] Geltung 2 schließt also das Trittbrettfahrerproblem nicht aus. Dies wurde bis vor kurzem in normativ orientierten Ansätzen zur Institutionalisierung weitgehend übersehen, vgl. neuerdings aber *Apel* (1992, S. 9), *Habermas* (1992).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Geltung 2b liegt vor, wenn die Wahl von Regeln zugleich die Standards des Handelns definiert. Sie berücksichtigt damit die intrinsische Motivation, die sonst in der Neoklassik nicht in die Analyse einbezogen wird, obwohl viele Autoren zugeben, daß es sie gibt (vgl. z.B. *Milgrom / Roberts* 1991, S. 42 ff.; *Kirchgässner* 1991, S. 61ff.).[32] In der ökonomischen Literatur diskutierte Beispiele für Normen der Geltung 2b sind Trinkgelder an Autobahnraststätten als Ausdruck der Zufriedenheit, der Gang zur Wahl aus Pflichbewußtsein („voter paradox") oder der Sachverhalt, daß Universitätsabsolventen für ein niedrigeres Gehalt zu arbeiten bereit sind, wenn sie sich mit dem Produkt identifizieren können, zu dessen Herstellung sie beitragen (vgl. *Frank* 1992, S. 169). Diese Regeln haben verpflichtenden Charakter. Dies bedeutet allerdings nicht, daß solche Regeln immer eingehalten werden. Schließlich gewinnt das Trittbrettfahrerproblem seine Relevanz gerade daraus, daß eigenes Trittbrettfahren um so lohnender ist, je mehr andere Individuen sich an die Regeln halten.[33] Entscheidend ist, daß Handeln nach solchen Regeln empirisch beobachtet werden kann, d.h., daß es ein Handeln gibt, das nicht auf bloße Routinen, Sanktionen oder Reziprozitätsnormen zurückgeführt werden kann.[34]

Auf den ersten Blick weist die Unterscheidung der verschiedenen Formen der Geltung eine Ähnlichkeit mit den Ebenen des individuellen Lernens auf. Der Unterschied besteht allerdings zum *einen* darin, daß Lernen (wie oben erwähnt, vgl. Abschnitt 2.2.4) ein irreversibler, fortschreitender Prozeß ist. Hingegen sind die verschiedenen Formen der Institutionalisierung wählbar, etwa in Form der Wahl einer sozialen Rolle. Zwischen den verschiedenen Formen gibt es vom Individuum kalkulierbare tradeoffs. Zum *anderen* läßt die Unterscheidung in verschiedene Formen der Geltung offen, wie die Normen *entstehen*. Jedoch gibt es innerhalb der Ökonomik einige Ansätze, welche versuchen, die Herkunft der Normen endogen zu erklären. Insoweit können diese Ansätze zugleich als Versuche interpretiert werden, Aspekte des organisationalen Lernens in die neoklassische Theorie zu intergrieren. Auf diese Ansätze wollen wir im folgenden eingehen.

---

32 Es mehren sich allerdings die Autoren, welche die Forderung aufstellen, die intrinsische Motivation sei explizit zu berücksichtigen (*Sen* 1987; *Elster* 1989; *Frey* 1990; *Frank* 1992).

33 Dies macht den Unterschied zwischen der hier vertretenen Position und der Position von Habermas (1981) aus. Die Begründung von Normen des Typs 2b garantiert noch nicht die Normbefolgung. Vielmehr hängt die Normbefolgung auch von den Kosten ab. Normen haben eine um so höhere Chance der Befolgung, je eher sog. „Kleinkostensituationen" (*Kirchgässner* 1991) gegeben sind. Zur empirischen Relvanz von Kleinkostensituationen im Bereich des Umweltverhaltens vgl. *Dieckmann / Preisendörfer* 1991.

34 Vgl. die Unterscheidung von Vertrauens- und Solidaritätsnormen bei *Vanberg* 1987. *Vertrauensnormen* setzen auf Reziprozität, d.h. auf langfristige Kooperationsvorteile, wie sie z.B. der Tit-for-Tat-Strategie (*Axelrod* 1991) zugrunde liegen. Sie können zur Bildung von kooperativen Mustern führen (vgl. *Schüßler* 1990 sowie Abschnitt 5 dieser Arbeit). Hingegen haben *Solidaritätsnormen* Kollektivgutcharater, d.h., sie gewähren keinerlei „selektive Anreize" (*Olson* 1968) zur Bildung kooperativer Muster.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

4. Institutionalisierung von individuellem Lernen aus der Sicht der ökonomischen Theorie

4.1 Vorbemerkung: Homo oeconomicus und Institutionalisierung

Die traditionelle ökonomische Theorie hat bisher *individuelles Lernen* nur als behavioristisches Lernen thematisiert. Dies ergibt sich zum einen daraus, daß sie mittels der Rationalitätsannahme Handeln auf Verhalten zurückführt (vgl. *Kirchgässner* 1991, S. 20). Zum anderen spielen in ihr Prozesse der Evaluation von Kontexten keine Rolle. Das Ergebnis ist das klassische SEU(subjectively expected utility)-Paradigma, welches grundlegend für das Konstrukt des homo oeconomicus ist: Dieser trifft eine rationale Wahl zwischen gegebenen Alternativen. Die Alternativen entsprechen als subjektive Repräsentationen der objektiven Wirklichkeit, d.h., sie sind bei allen Individuen und zu allen Zeiten nur von der objektiven Situation abhängig. Obgleich diese sog. Invarianzannahme seit dem experimentellen Nachweis von Entscheidungsanomalien als gescheitert gelten kann (vgl. oben Abschnitt 2.2.3 zu Tversky / Kahneman), wird sie von vielen Ökonomen nach wie vor zugrunde gelegt. Für dieses Vorgehen werden mehrere Begründungen angeführt (vgl. *Frey* 1990, S. 168 ff.; *Kirchgässner* 1991, S. 152 ff.): Beispielsweise würden sich Anomalien durch den Aggregationsprozeß ausgleichen oder durch den Wettbewerb verschwinden. Selbst wenn man diesen Begründungen auf der Makroebene folgen würde (vgl. zu den Einwendungen *Frey* 1990, S. 174 ff.), so sind sie auf der Mikroebene nicht stichhaltig, weil es hier darum geht, wie Individuen in Organisationen zu objektiven Repräsentationen finden, d.h. ihre bestehenden „frames" kritisch evaluieren.

Im Ergebnis führt dies dazu, daß — wie in der Einleitung bereits festgestellt — die neoklassiche Theorie Lernen nur als den Erwerb von Informationen, nicht aber als die Generierung von Wissen erklärt. Dies würde nämlich die Einbeziehung von unterschiedlichen kontextspezifischen „frames" bzw. die kritische Evaluation dieser „frames" voraussetzen. Auf diese Weise werden in der Neoklassik sogar Innovationen als Maximierung über ein gegebenes Set von Alternativen behandelt, wohingegen die Herkunft der Alternativen unbeleuchtet bleibt. Zu fragen ist jedoch: „Does the assumption that ‚actors maximize' help one to analyze situations where some actors are not even aware of a possibility being considered by others?" (*Nelson* 1991, S. 66). Die Annahme hilft kaum in der eingangs beschriebenen aktuellen Wettbewerbssituation. Sie ist nur in standardisierten Situationen mit stabilen Rahmenbedingungen realistisch. Entscheidungen von strategischer Relevanz weisen diese Eigenschaft allerdings nicht auf (vgl. *Steinmann / Schreyögg* 1986).

In bezug auf das Problem der *Institutionalisierung* ist festzustellen, daß die neoklassische Theorie bisher meistens nur Institutionalisierung vom Typ Geltung 1 behandelt hat. So wird beispielsweise versucht, die Existenz von Unternehmens-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

kulturen auf die Wirkung von Sanktionen zurückzuführen (vgl. z.B. *Kreps* 1990; *Casson* 1991; *Föhr / Lenz* 1992).[35] Jedoch gibt es hier einige interessante Varianten, die für eine Theorie des organisationalen Lernens deshalb fruchtbar gemacht werden können, weil sie in unterschiedlicher Weise Elemente des Wissenserwerbs (im Unterschied zum Informationserwerb) als exogene Faktoren zulassen. Es sind dies zugleich Konzepte, welche zur Frage der Interdisziplinarität Stellung nehmen und in diesem Zusammenhang das Verhältnis von theoriegeleitetem Wissen und verschiedenen Formen der Institutionalisierung thematisieren.

4.2 Der homo oeconomicus als analytische Kunstfigur (Parallelmodell)

Eine erste Variante relativiert das neoklassische Modell. Dieses Modell, das von Homann und Koautoren vertreten wird, plädiert nicht nur dafür, daß alternative Sichtweisen eine Berechtigung haben. Vielmehr sollen zwischen dem neoklassischen und anderen Modellen „Paralleldiskurse" geführt werden. Die Autoren gehen davon aus, daß der homo oeconomicus nicht mit dem Anspruch auftritt, eine zutreffende Beschreibung der Realität darzustellen; vielmehr stelle er ein heuristisches Schema oder eine Suchanweisung dar, welche die ökonomische von anderen Perspektiven unterscheidet. Dieses Schema lautet: „Erkläre Handeln so, als ob die Handelnden ihren erwarteten Nutzen unter Restriktionen maximieren würden" (*Homann / Suchanek* 1989). Der homo oeconomicus wäre damit kein empirisches Erklärungsmuster, sondern ein Konstrukt, das „on rather analytical than empirical grounds" (*Brennan / Buchanan* 1985, S. 59) gewählt wird. Der Vorwurf, es handle sich dabei um ein verkürztes Menschenbild, sei eine Tautologie. Ebenso sei seine Verteidigung mit dem Argument des Realismus lediglich eine „Prolongierung einer systematisch verfehlten Auffassung von einzelwissenschaftlicher Forschung" (*Homann* 1988, S. 117), welche zu interdisziplinärem Dilettantismus führe. Der homo oeconomicus sei weder richtig noch falsch, sondern erweise sich in bestimmten Forschungszusammenhängen als zweckmäßig oder nicht.[36]

Allerdings weise diese Kunstfigur eine beträchtliche Leistungsfähigkeit auf. Entgegen allen Versuchen, den homo oeconomicus zu erweitern, habe die Ökonomik ihren Aufschwung in den letzten Jahrzehnten nicht ihren Erweiterungen, sondern gerade ihrer Radikalisierung zu verdanken, wofür insbesondere der Name Gary S. Becker stehe. Die Zahl der fruchtbaren Anwendungen des homo oeconomicus habe sich sprunghaft ausgeweitet. Der Grund liege darin, daß in modernen

---

35 Zur Kritik vgl. *Löhr / Osterloh* (1993); *Osterloh* (1994).

36 Die Autoren geben allerdings nicht an, welche Kriterien dabei zugrunde gelegt werden sollen. Doch macht es wenig Sinn, hierbei wieder ausschließlich auf disziplinäre Sichtweisen abzustellen, weil dies eine Immunisierung gegenüber empirischer Überprüfung bedeutet. Disziplinäre Perspektiven kann man nicht falsifizieren (vgl. *Homann / Suchanek* 1989, S. 82). Es bleiben also nur die Kriterien der aufgeklärten Praxis.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Gesellschaften der traditionelle Konsens über Normen endgültig zerbrochen sei. Verhaltenskoordinierung könne nur noch über Ergebnisse erfolgen (vgl. *Homann / Pies* 1994, Ziffer 27). Institutionelle Arrangements sollten daher einem H-O-Test (homo oeconomicus-Test) unterworfen werden, der so ausgestaltet sein sollte, daß unerwünschte Handlungen destabilisiert, erwünschte Handlungen hingegen stabilisiert werden.

Die Frage, wie dies auf der Ebene psychischer Abläufe beschrieben werden kann, wird explizit nicht gestellt. Erkenntnisse der Psychologie sollen nur in reduzierter Form unter Kategorien wie Einkommen oder Prestige subsumiert werden. Wer dabei die Frage nach den *Ursachen* für die Orientierung an der Anreizstruktur stelle, mithin die Prozesse des Erkenntnisgewinns analysieren wolle (wie z.B. die Evolutionsökonomie), steige aus dem ökonomischen Forschungsprogramm aus. „Statt dessen wird in einer Als-ob-Konstruktion den Individuen das Wissen zugeschrieben, das diese möglicherweise erst am Ende des Prozesses haben" (*Suchanek* 1991, S. 85). Damit wird das klassische SEU-Paradigma (zumindest als analytische Fiktion) beibehalten.

Die Kunstfigur des homo oeconomicus setzt damit von vornherein standardisierte Situationen mit definierten Kosten-Nutzen-Relationen voraus. In solchen Situationen kann angenommen werden, daß von Eigeninteresse geleitete Individuen auf entsprechende Anreize ähnlich reagieren, mithin Mechanismen der Institutionalisierung durch Sanktionen greifen (vgl. *Homann / Suchanek* 1992, S. 15).

Damit bleiben Homann und Koautoren auf der Ebene der *Wissensdimension* ganz bewußt im Rahmen der Invarianzannahme der klassischen SEU-Theorie. In bezug auf die *Institutionalisierungsdimension* stellen sie ausschließlich auf Geltung 1 ab, d.h. Institutionalisierung durch Sanktionen. Dies ergibt sich aus der Behandlung des homo oeconomicus als analytische Kunstfigur. Zwar könne man andere Formen des Lernens in das ökonomische Modell übersetzen, beispielsweise wenn *North* (1988, S. 88) Ideologien als eine kognitive Sparmaßnahme kennzeichnet. Umgekehrt sei es sinnvoll, ökonomische Sachverhalte in terms anderer Disziplinen zu übersetzen. Wissenschaftlicher Fortschritt allerdings sei nur möglich, wenn jedes wissenschaftliche Paradigma empirische Erscheinungen in *seine eigene Sprache* übersetze, um anschließend mit den anderen Disziplinen Paralleldiskurse zu führen. Insofern verläßt zwar jeder das ökonomische Forschungsprogramm, der die Frage nach organisationalem Lernen im Sinne eines Erwerbs von theoretischem Wissen stellt. Ökonomikextern sei es aber durchaus sinnvoll, diese Fragen zu behandeln.

Die geschilderte Position wirft drei Probleme auf: *Erstens* ist es offensichtlich, daß das Phänomen des Wissenserwerbs und damit der gesamte Bereich der Innovationsforschung zum exterritorialen Gebiet der neoklassischen Ökonomik erklärt wird. Mit der Übersetzung der Wissensdimension in „terms of economics" werden nämlich zugleich offene Probleme (d.h. Probleme, die durch fundamenta-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

le oder strukturelle Unsicherheit gekennzeichnet sind) immer in geschlossene überführt (d.h. in Probleme, bei denen der mögliche Alternativenraum schon definiert ist). Dies folgt zwingend aus der Behandlung des homo oeconomicus als analytischer Kunstfigur. Die Neoklassik schließt sich damit von einem der wichtigsten Themenbereiche moderner Wirtschaften aus und büßt damit praktische Relevanz ein.

Zweitens bleibt völlig offen, wie genau der sogenannte „Paralleldiskurs" geführt werden soll bzw. wie der Praktiker mit den verschiedenen disziplinären Sichten umgehen soll. Seine Probleme halten sich nämlich nicht an die von den Wissenschaften definierten Grenzen, vielmehr gilt: „ Die Paxis ist a-disziplinär" (Hans *Ulrich* 1988, S. 177). Für die *problemorientierte,* nicht *disziplinorientierte Lösung* praktischer und theoretischer Fragen enthält die Position von Homann keinerlei Hilfestellung.

Drittens halten die Autoren ihren Anspruch selber nicht durch, den homo oeconomicus nur als ein heuristisches Konstrukt einzusetzen. Wenn sie empfehlen, institutionelle Arrangements einem H-O-Test (homo-oeconomicus-Test) zu unterwerfen, sind sie schon dabei, die analytische Kunstfigur (kryptonormativ) in die Praxis umzusetzen — mit den oben geschilderten Problemen der Übersetzung von offenen in geschlossene Probleme für die Innovationsforschung.

Trotz der geschilderten Probleme bietet die Position von Homann und Koautoren ein außerordentlich hilfreiches analytisches Raster an, auf dessen Hintergrund die folgenden Ansätze besser verständlich werden.

4.3 Der homo oeconomicus als imperialistisches Konstrukt (Dominanzmodell)

Esser (1990, 1991) und *Lindenberg* (1990) treten mit einem viel radikaleren Anspruch als das Modell von Homann auf. Sie erweitern den Anspruch des ökonomischen Forschungsprogramms grundsätzlich: Die Ökonomik dominiere alle anderen Programme, weil man alle anderen Programme *restlos* in das ökonomische Programm übersetzen könne. Der umgekehrte Weg sei hingegen von vornherein zum Scheitern verurteilt (vgl. *Lindenberg* 1990, S. 251).

In unserem Zusammenhang ist von Bedeutung, daß die Autoren zwar einerseits Informationen als Daten behandeln und den Lernprozeß als Tauschprozeß und nicht als Produktionsprozeß modellieren: Lernen wird als ökonomisch motivierte rationale Entscheidung in Knappheitssituationen interpretiert. Knapp ist die menschliche Informationsverarbeitungskapazität, und entschieden wird zwischen verschiedenen kognitiven Schemata, Skripts, Faustregeln, Rezepten oder Routinen, die vorgegeben sind. Wie man sieht, wird hierbei der theoretische Reflexionsprozeß bei der Bildung kognitiver Strukturen auf Prozesse der Institutionalisierung zurückgeführt, genauer auf Institutionalisierung vom Typ Geltung 2a.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Andererseits geben sie aber — im Unterschied zu Homann — die Invarianzannahme des klassischen SEU-Paradigmas auf, führen dies allerdings nicht auf den Erwerb von theoretischem Wissen, sondern auf Institutionalisierungsmechanismen zurück.

Esser (1990, 1991) stellt hierfür ein zweistufiges Selektionsmodell vor. Auf der *ersten Stufe* wird ein durch die Situation induziertes Rezept oder eine Routine ohne nähere Nachprüfung angewendet. Solche Rezepte sind kognitiv leicht verfügbar, d.h., die Informationskosten sind niedrig. Weil sie meist in einem institutionellen Zusammenhang zur Verfügung gestellt werden, geben sie zugleich eine hohe Sicherheit und eine Verläßlichkeit in bezug auf die Abstimmung mit anderen Personen. Sie stellen damit eine „kognitive Sparmaßnahme" dar (*North* 1988, S. 50) und bewirken zugleich eine deutliche Einsparung von Transaktionskosten (vgl. *Esser* 1990, S. 235). Erst wenn die gewohnten Routinen nicht mehr den Erwartungen entsprechen, werden auf der *zweiten Stufe* neue Routinen gesucht. Eine Änderung von verfestigten Routinen wird immer dann angestrebt, wenn der erwartete Nutzen aus der Anwendung neuer Routinen unter der Berücksichtigung der Wahrscheinlichkeit, neue Routinen zu finden und zu implementieren, größer ist als die Suchkosten für diese Routinen. Die Wahl der Routinen hängt dabei auch vom Erinnerungsvermögen des Individuums ab. Aber wie wird gewährleistet, daß die neuen Routinen der Akteure mit den „objektiven" Relevanzstrukturen übereinstimmen?

Zur Beantwortung dieser Frage hat *Lindenberg* (1990) ein sog. „Diskriminations-Modell" entwickelt. Auch dieses geht von der beschränkten Informationsverarbeitungskapazität aus. In einer *ersten Phase* werden Alternativen und Resultate pro Alternative unterschieden und nach den Präferenzen des Individuums geordnet. Dabei hat das Individuum immer spezifische situationale Ziele im Kopf, z.B. „ein schlauer Konsument sein" oder „ein guter Freund sein". Es kann wegen seiner kognitiven Beschränkungen immer nur ein Ziel aktualisieren. In einer *zweiten Phase* bewertet das Individuum die Alternativen mit der subjektiven Auftretenswahrscheinlichkeit. In der *dritten Phase* wählt das Individuum nicht einfach die Alternative mit dem höchsten Erwartungswert, sondern jede Alternative nur mit einer gewissen Wahrscheinlichkeit. Diese hängt von zwei Faktoren ab: erstens von der Differenz der Nutzenerwartung für die betrachteten Alternativen, zweitens von der Bedeutsamkeit („salience") des Ziels. Die Bedeutsamkeit ist ein Maß dafür, wie dominant das Ziel in der gegebenen Situation ist. In einer *vierten Phase* wird die Bedeutsamkeit des Ziels selbst evaluiert. Es geht dabei um dessen indirekte Verbindung zu anderen Lebensbereichen, beispielsweise wenn der Erwerb von Reputation in einem Bereich auch in anderen Bereichen nützt.

Im Unterschied zu Esser geht Lindenberg auf die Entstehung der Ziele ein: Sie sind kulturell bedingt, wobei Kultur die für eine geschichtliche Situation charakteristischen sozialen Produktionsfunktionen umfaßt. Produktionsfunktio-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

nen geben die Herstellbedingungen für diejenigen Güter an, die allgemein angestrebt werden, nämlich insbesondere physisches Wohlbefinden und soziale Wertschätzung.

Das Modell von Esser und Lindenberg ist damit zwar in der Lage, unterschiedliche situationale Ziele oder Rollenvorstellungen zu berücksichtigen und damit das klassische SEU-Paradigma zu überwinden. Sie führen diese Erweiterung ausschließlich auf *Institutionalisierungsmechanismen* zurück, nicht auf theoriegeleitetes *Wissen*. In bezug auf die Institutionalisierungsmechanismen erweitern sie allerdings das Repertoire gegenüber dem Ansatz von Homann: Neben der Institutionalisierung durch Sanktionen (Geltung vom Typ 1) wird auch Institutionalisierung durch die Wahl von Routinen (Geltung vom Typ 2a) einbezogen. Mit dieser Erweiterung wird zwar erreicht, daß in das Tauschmodell der Neoklassik nunmehr auch unterschiedliche Routinen, Rollenschemata oder Faustregeln eingehen. Mit der Einbeziehung der Kultur als sozialer Produktionsfunktion wird sogar eine Erklärung dafür geboten, wie solche Routinen entstehen. Dennoch bleiben zwei Probleme ungelöst.

Zum *ersten* wird nicht erklärt, warum sich kollektive Entscheidungsanomalien wie „groupthink" (vgl. oben) herausbilden, über längere Zeit halten und sogar als „pathologische Unternehmenskulturen" (vgl. *Kets de Vries / Miller* 1984) institutionalisieren können *(Wissensdimension)*. Der Grund für dieses Defizit liegt darin, daß Prozesse der Genese von neuen Routinen von den Autoren nicht in ihre Analyse einbezogen werden (vgl. *Srubar* 1992). Die eigentliche Pointe des Double-loop-Lernens, nämlich wie neue (und manchmal auch falsche) organisatorische Routinen entwickelt werden, bleibt außer Betracht. Prozesse der Orientierung in Situationen fundamentaler Verunsicherung (vgl. *Siegenthaler* 1993) oder struktureller Unsicherheit (vgl. *Minkler* 1993) werden nicht behandelt. Solche Verunsicherungen treten jedoch regelmäßig auf, wenn verschiedene „frames" zusammentreffen (z.B. im interkulturellen Management, beim Zusammentreffen verschiedener Unternehmenskulturen oder auch beim Zusammenprall verschiedener disziplinärer Sichtweisen). Gerade dies macht den Kern des organisationalen Lernens und der Innovation aus. In diesen Situationen kann man nicht auf vorgefertigte Routinen zurückgreifen, sondern man muß sich zuerst auf neue, gemeinsame Alternativensets verständigen, *bevor* diese institutionalisiert werden können.

Daraus resultiert das *zweite* Problem, welches darin begründet liegt, daß die Autoren wichtige Aspekte der Institutionalisierung ausklammern *(Institutionalisierungsaspekt)*. Bei den Verständigungsprozessen über gemeinsame Alternativensets sind zwei Problemklassen zu unterscheiden (vgl. *Vanberg* 1984), welche zugleich die zwei Funktionen von Kultur kennzeichnen (vgl. *Osterloh* 1994). Es ist dies zum einen das Problem der *Koordination*, d.h. das Problem der Erzeugung von Verhaltensregelmäßigkeiten in Situationen, in denen kein Interessenkonflikt auftritt. In solchen Fällen ist es den Individuen gleichgültig, welche Regelung gewählt wird, solange alle dieselbe Wahl treffen (z.B. ob rechts oder links

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

gefahren wird).[37] Diese Probleme der Verhaltensabstimmung werfen keine größeren Schwierigkeiten auf, sie lassen sich durch wechselseitige Information oder sogar im Wege der „tacit coordination" durch „focal points" lösen (vgl. *Schelling* 1960) und mit dem klassischen transaktionskostentheoretischen Instrumentarium analysieren.[38]

Zum anderen handelt es sich um das Problem der *Kooperation,* welches viel schwerer lösbar ist. Bei diesem treten zwischen den Individuen Interessenkonflikte bei der Verhaltenskoordinierung auf, die auf Ausbeutungsmöglichkeiten zurückgehen. Sie werden üblicherweise als Gefangenendilemmata beschrieben. Diese Probleme stellen, wie die reichhaltige Literatur über das Gefangenendilemma beweist, in diesem Zusammenhang die derzeit wichtigste Herausforderung des ökonomischen Ansatzes dar. Sie werden im Rahmen spieltheoretischer Ansätze behandelt (vgl. unten). Weder Esser noch Lindenberg gehen auf diese Schwierigkeiten ein.[39]

Zusammenfassend kann zum Modell von Esser und Lindenberg festgestellt werden, daß sie auch mit der Erweiterung des SEU-Modells um die Einbeziehung von verschiedenen Routinen die eigentliche Pointe von Lernen und Innovationen in Organisationen verfehlen. Sowohl die Entstehung von neuen Routinen *(Wissensaspekt)* wie deren Institutionalisierung im Falle von Interessendivergenzen *(Institutionalisierungsaspekt)* wird in ihrem Modell nicht erklärt.

4.4 Der homo oeconomicus als ergänzungsbedürftiges Konstrukt (Sequenzmodell)

Eine dritte Variante innerhalb der ökonomischen Theorie nimmt eine Trennung zwischen dem *Wissensaspekt* und dem *Institutionalisierungsaspekt* vor und weist beide Aspekte jeweils getrennten Disziplinen zu. Dieses Modell wurde von B. S. Frey entwickelt (vgl. *Frey.* 1990, 1992). Frey trägt damit den bisher aufgezeigten Defiziten des neoklassischen Modells Rechnung, indem er von vornherein von dessen Ergänzungsbedürftigkeit ausgeht. Dies findet seinen Ausdruck darin, daß zwei Phasen der Modellbildung vorgeschlagen werden:

In der *ersten Phase* geht es um die Entwicklung und Definition der für das Individuum relevanten Alternativen. Für diese ist die Psychologie zuständig. In der *zweiten Phase* geht es um die Auswahl zwischen diesen Alternativen. Dies

---

[37] Diese stellen auf Erwartungserwartungen ab vom Typ „Was würde ich tun, wenn ich er wäre, und er überlegen würde, was ich tun würde" (vgl. *Schelling* 1960, S. 54).

[38] Vgl. z.B. das Konzept der „scale returns of communication" von *Arrow* (1991).

[39] Hingegen weist *North* (1988, S. 56 ff.) darauf hin, daß die Ökonomik eine Theorie der Ideologie (im Sinne von geteilten mentalen Modellen) und der Institutionalisierung von Ideologie benötigt, um die Probleme der Kooperation berücksichtigen zu können.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ist Sache der Ökonomik. Dieses Zweiphasenmodell demonstriert Frey anhand von zwei Problemkreisen.

(1) Der erste betrifft das *ipsative Modell* des menschlichen Verhaltens. Hierbei geht Frey nicht nur von der Erkenntnis aus, daß sich die Invarianzannahme der klassischen SEU-Theorie nicht mehr aufrechterhalten läßt, sondern daß alternative kognitive frames durchaus in verschiedener Weise für das Handeln der Individuen relevant sind. Deshalb unterscheidet er zwischen einem objektiven und einem ipsativen Möglichkeitsraum. Der ipsative Möglichkeitsraum ist der von einer bestimmten Person als für sich selbst für relevant gehaltene Möglichkeitsraum (z. B. wenn Elizabeth Taylor bei jeder ihrer zahlreichen Hochzeiten überzeugt verkündet, daß diese die letzte sei). Frey schließt nicht aus, daß Individuen ihren ipsativen Möglichkeitsraum auch strategisch ausdehnen oder reduzieren können, nämlich durch strategische Selbstbindungen (z. B. Auferlegung einer Verlobungszeit). Insofern wird die Trennung der „psychologischen" von der „ökonomischen" Phase nicht strikt durchgeführt.

(2) Der zweite Problemkreis betrifft die Einbeziehung der *intrinsischen* neben der extrinsischen Motivation. Die intrinsische Motivation kann erklären, warum Individuen in Situationen des Interessenkonfliktes freiwillig zu Kooperation bereit sind. Sie führt zu Institutionalisierungsmechanismen vom Typ Geltung 2b. Intrinsische Motivation kann allerdings durch extrinsische Motivation zerstört werden, nämlich dann, wenn einer Person eine Belohnung für eine Aktivität angeboten wird, die sie gemäß ihren Präferenzen ohnehin unternehmen würde. Die intrinsische Motivation droht dann verlorenzugehen, weil sie funktionslos geworden ist (vgl. *Frey* 1990, S. 157; 1994). Es geht dann darum, durch Maßnahmen der Institutionalisierung zu versuchen, ein optimales Verhältnis von intrinsischer und extrinsischer Motivation zu erzielen. Auch hier wird die Trennung der „psychologischen" von der „ökonomischen" Phase nicht strikt durchgeführt.[40]

Das Modell von Frey weist sowohl in bezug auf den *Wissensaspekt* als auch in bezug auf den *Institutionalisierungsaspekt* Vorteile für die Erfassung des organisationalen Lernens auf, ohne daß Frey selbst diese Fragestellung aufgreift.

Die Sequentialisierung bewirkt, daß das Modell auch jenseits des Gedankens des ipsativen Verhaltens prinzipiell offen ist für den *Wissensaspekt,* d.h. für die (alltags-)theoriegeleitete Generierung von neuen Alternativen, die das Wesen der Innovation ausmachen. Hinsichtlich des *Institutionalisierungsaspektes* liegt der Vorteil darin, daß empirisch ermittelt werden kann, in welchen Situationen die intrinsische oder die extrinsische Motivation für die Erklärung menschlichen Handelns zutrifft. Damit können empirisch gesicherte Aussagen darüber gemacht werden, wann und unter welchen Bedingungen Geltung vom Typ 2b auftritt.

Allerdings ist bisher noch keine theoretische Integration des homo oeconomicus und des homo psychologicus gelungen.[41] Ebensowenig wie bei Esser und Lindenberg wird modellimmanent erklärt, wie trotz konfligierender Handlungs-

---

[40] Gleichzeitig wird aber der Anspruch auf die Dominanz des ökonomischen Modells beibehalten.

[41] Zu einem neueren Versuch der Integration vgl. *Eichenberger / Frey* (1993).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

muster Institutionen möglich sind.[42] Die Folge davon ist, daß Institutionen funktionalistisch eingeführt werden: Sie existieren, weil sie gebraucht werden. Damit handelt es sich streng genommen um einen funktionalistischen Fehlschluß, der innerhalb einer Theorie nicht befriedigen kann, die Institutionen individualistisch erklären will (vgl. *Opp* 1983, S. 96).

5. Die Institutionalisierung von individuellem Lernen aus der Sicht der Spieltheorie

5.1 Zur Grundidee der Spieltheorie

Wir haben gesehen, daß das, was das organisationale Lernen vom individuellen Lernen unterscheidet — nämlich die Aspekte der *Institutionalisierung* und der *Kooperation* — in den bisher behandelten ökonomischen Modellen nicht endogen erklärt werden konnte. Insbesondere wird nicht erklärt, wie ohne Inanspruchnahme funktionalistischer Erklärungsweisen oder intrinsisch motivierter Geltungsansprüche Institutionen zustande kommen. Im folgenden wollen wir erörtern, ob und inwieweit die non-kooperative Spieltheorie hier weiterführt. Zur *Wissensdimension* trägt die Spieltheorie dagegen derzeit wenig bei, weil sie im Normalfall von einer strengen Rationalitätsdefinition ausgeht.[43]

Die Spieltheorie fragt danach, welche Handlungsstrategien von Individuen in einer sozialen Situation (in einem Spiel) zu stabilen Interaktionen (Gleichgewichten) führen, und zwar u. a. in den Situationen, in denen Konflikte zwischen den Interaktionspartnern auftreten. Sie sucht aufzuzeigen, wie diese Konflikte unter rationalen Egoisten bewältigt werden können. Wenn es ihr gelänge, die Bedingungen für eine stabile Kooperation aufzuzeigen, hätte sie das Institutionalisierungsproblem auch für den Fall konfliktärer Interessen gelöst.[44]

Die Spieltheorie verbindet die Perspektive der Maximierung des Erwartungsnutzens mit dem Aspekt der strategischen Interaktion.

Sie legt ihren Aussagen im wesentlichen folgende Annahmen zugrunde:

(1) Die Individuen werden als strategische Erwartungsnutzen-Maximierer modelliert.

---

[42] Allerdings führt Frey neuerdings Experimente zum Gefangenendilemma durch, die aufzeigen sollen, bei welchen Bedingungen kooperative bzw. defektive Lösungen zu erwarten sind.

[43] Neuerdings wird versucht, den Aspekt der bounded rationality in die Spieltheorie einzubeziehen (vgl. *Fudenberg / Kreps* 1992; *Fudenberg / Levine* 1993).

[44] Zur Erinnerung: Es geht um die Auswahl zwischen verschiedenen konkurrierenden organisatorischen Routinen. Für dieses Problem haben, wie oben gezeigt, Esser, Lindenberg und Frey keine Lösung aufgezeigt.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

(2) Die Individuen berücksichtigen in ihren Entscheidungen die möglichen Entscheidungen und Handlungen aller anderen Aktoren.

(3) Die Individuen nehmen teil an einer sozialen Situation (bzw. einem Spiel), die durch eine feste Struktur beschrieben werden kann.

(4) Die Individuen kommunizieren nicht im Spiel. Symbolische und verbale Kommunikation ist ausgeschlossen. Es ist allenfalls „cheap talk" möglich (vgl. aber hierzu *Kreps* 1991; *Johnson* 1993).

Mit der letzten Annahme will die Spieltheorie — im Gegensatz zur Argumentation in der Literatur zum organisationalen Lernen — aufzeigen, wie Institutionen ohne Kommunikation zustande kommen können. Wir beschränken uns in unserer Argumentation auf Spiele, die folgende Charakteristiken aufweisen:[45]

(1) Die Individuen können unvollständige Informationen über die Handlungen der anderen Individuen im Spiel haben.

(2) Das Spiel ist ein „repeated game", d.h., die Individuen spielen in jeder Periode dasselbe Spiel. Dies schließt ein, daß es keine Möglichkeiten für Investitionen (z.B. in „human capital") geben kann.

(3) Im Spiel herrscht „perfect recall": Einmal verarbeitete Information wird nicht vergessen.

Aus der allgemeinen Fragestellung der Spieltheorie leiten wir für unsere weitere Argumentation folgende Teilfragen her:

(1) Welche Bedingungen müssen gelten, daß wir die Etablierung eines Gleichgewichtes für eine gegebene Spielstruktur bestimmen können? Die Beantwortung dieser Frage zeigt, wie sich Unterschiede in der Spielstruktur auf die Veränderung der Gleichgewichte auswirken.

(2) Kann die Entstehung systematischer neuer Verhaltensmuster im Rahmen eines vorgegebenen Spiels beschrieben werden? Dieses würde uns Einsichten in den Institutionalisierungsprozeß ermöglichen.

(3) Gibt es eine Möglichkeit, trotz der Annahme der vollständigen Rationalität Lernprozesse im Spiel zu modellieren (vgl. dazu auch *Kreps* 1991; *Rubinstein* 1991)?

5.2 Folk Theorem und die Folgen

Die Spieltheorie legt ihrer Argumentation ein anderes Gleichgewichtskonzept zugrunde als die neoklassische Ökonomie.[46] Dabei gibt es verschiedene spieltheo-

[45] Deren Plausibilität bzw. Problematik wird sich in der weiteren Argumentation zeigen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

retische Lösungskonzepte, deren gemeinsame Idee für die Identifikation von Gleichgewichtssituationen die folgende ist: Jeder Spieler definiert seine Strategien[47] so, daß er für jede mögliche Entscheidung aller anderen Aktoren seine Entscheidungen bestimmt. Im Gleichgewicht kann kein Aktor durch die Veränderung seiner Entscheidungen gewinnen, gegeben die möglichen Entscheidungen aller anderen Aktoren.

Auf diesem Hintergrund kann die Frage diskutiert werden, ob Kooperation möglich ist: Ist die Institutionalisierung kooperativer Verhaltensmuster ohne Kommunikation und ohne Einbeziehung sozialer Bedürfnisse in die Nutzenfunktion der Individuen eine stabile Gleichgewichtssituation?[48]

In „repeated games" mit unendlich vielen Perioden und / oder mit unvollständiger Information ist allerdings dann Kooperation nur eine mögliche Strategie. Auch die Nichtkooperation ist eine mögliche Lösung des Spiels. Dieses Ergebnis folgt aus dem Folk Theorem (*Taylor* 1976): Danach kann sich *jede* mögliche Lösung als Gleichgewicht etablieren, wenn die Spieler genügend geduldig sind.

Unter welchen Bedingungen etabliert sich aber aus den vielen möglichen Gleichgewichten die Kooperation als stabile Lösung? Die Spieltheorie beantwortet diese Frage, indem sie in die allgemeine Spielstruktur zusätzliche Annahmen über die Aktoren bzw. über das Spiel einführt. Diese Annahmen werden in der Literatur als „refinements" diskutiert.

Für Spiele mit *vollständiger Information* sind unter anderem die folgenden „refinements" relevant:

(1) Durch die Einführung *zusätzlicher* Informationen (die vorher durch die formale Modellierung wegdefiniert wurden) können einzelne Lösungen favorisiert werden. Prominent geworden ist die Idee der „focal points" (*Schelling* 1960), auf die bereits hingewiesen wurde: Die Spieler verfügen über ein

---

46 Allerdings kann das neoklassische Gleichgewicht als ein Spezialfall des spieltheoretischen Gleichgewichts interpretiert werden, nämlich für den Fall, für den die Entscheidungen aller anderen Aktoren nur als Preissignale wahrgenommen werden.

47 Der spieltheoretische Strategiebegriff ist wesentlich enger als der Strategiebegriff der Managementlehre. Dieser umfaßt zwar wie in der Spieltheorie die systematische Einbeziehung von Reaktionen der anderen Aktoren in die eigene Entscheidungsfindung, allerdings werden dabei Prozesse der materiellen und kognitiven Kontextbeeinflussung (z.B. durch „framing") miteinbezogen.

48 In den Sozialwissenschaften wird seit langem das Hobbes'sche Problem diskutiert: Wie ist soziale Ordnung ohne institutionalisierte Vorgaben möglich? *Kliemt* (1991) unterscheidet dabei zwei Aspekte: Auf der einen Seite kann man fragen, unter welchen Bedingungen in einem bestehenden Rahmen neue Regeln einer sozialen Ordnung entstehen können (Hobbes I). Auf der anderen Seite besteht die viel fundamentalere Frage, wie in einer Anarchie grundsätzlich überhaupt soziale Regeln entstehen können, ohne daß bereits ein Rahmen vorausgesetzt wird (Hobbes II). Die oben behandelte Frage der Spieltheorie behandelt nur Hobbes I. Damit werden zugleich die Grenzen der spieltheoretischen Argumentation deutlich.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

gemeinsames soziales Wissen, etwa aufgrund ihres kulturellen Hintergrundes, welches gewisse Lösungen eher wahrscheinlich macht.[49]

(2) Durch die Einführung des Kriteriums der *Robustheit* verändert sich die Plausibilität einiger Lösungen. Es setzen sich langfristig die Strategien als Lösungen durch, die trotz leicht veränderter Spiele zu Gleichgewichten führen und damit gegenüber diesen Veränderungen robust sind.[50]

(3) Durch die Einführung der Möglichkeit, daß einzelne Spieler *Fehler* machen, müssen die Mitspieler diese Fehlermöglichkeit in ihre Entscheidungen miteinbeziehen. Sie werden deshalb gewisse Strategien und damit Lösungen vorziehen (*Fudenberg / Tirole* 1991).

In Spielen mit *unvollständiger Information* sind weitere Variationen möglich:

(4) Durch *Pre-play-Kommunikation* können die Spieler vor dem Spiel gemeinsame Konventionen herausbilden (analog zu den „focal points"), verbunden mit Signalen bzw. Handlungsweisen, über die sie den anderen Aktoren die Einhaltung der Konventionen im Spiel anzeigen. Ähnlich wie gemeinsame traditionelle oder kulturelle Muster erlauben diese Konventionen eine Reduktion der möglichen Lösungen.

(5) Im Konzept des *„common knowledge"* (vgl. Schelling 1960; *Fudenberg / Tirole* 1991) wird zugelassen, daß der einzelne Aktor weiß, daß die anderen Individuen bestimmte Informationen auch haben. Dieses Wissen erlaubt es den Aktoren, aus den Entscheidungen und Handlungen der Individuen auf weiteres, nicht gemeinsames Wissen zu schließen.

Die verschiedenen „refinements" reduzieren das Problem der multiplen Gleichgewichte, indem sie mittels Zusatzeigenschaften die möglichen Lösungen des Modells einschränken. Damit führen sie Eigenschaften wieder ein, die in der Modellierung mit dem Ziel formaler Strenge ausgeschlossen wurden (zur Problematik vgl. auch *Fudenberg / Tirole* 1987). Dabei wird ein Teil der Bedingungen wieder exogenisiert, beispielsweise die Frage, wie „focal points" möglich sind. Darüber hinaus bleibt die Frage offen, wie rationale Aktoren überaupt Fehler machen können (vgl. Punkt (3) oben).

---

[49] Es ist auf dem Hintergrund der Transaktionskostentheorie sogar möglich, die Funktion von „focal points" als ökonomisch sinnvoll zu begründen: Es bedeutet eine Reduktion der Transaktionskosten, wenn die Individuen sich aufgrund gemeinsamer „frames" in sozialen Interaktionen auf stabile Gleichgewichte einstellen bzw. diese für ihre Entscheidungen als plausibel voraussetzen können. Die „focal points" müssen in dieser Argumentation allerdings wiederum funktionalistisch eingeführt werden.

[50] Robustheit zeichnet auch gemeinsame kulturelle „frames" aus. Stabile „frames" haben die Eigenschaft, daß sie sich in den verschiedensten Entscheidungssituationen bewähren und somit zu stabilen Verhaltensmustern werden können. Für die Aktoren bedeutet die Existenz von „frames" Geltung vom Typ 2a.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Ungeachtet dieser Einschränkungen kann man von der Klarheit der Unterscheidungen profitieren, welche die Spieltheorie liefert, und diese Unterscheidungen auf das Konzept des organisationalen Lernens anwenden.

5.3 Lernen im spieltheoretischen Modell

5.3.1 Individuelles Lernen (Wissensdimension)

Wie bereits erwähnt schließt die der Spieltheorie zugrunde gelegte Annahme der vollständigen Information Prozesse des Lernens aus. Neuere Entwicklungen in der Spieltheorie versuchen allerdings, in Situationen der unvollständigen Information fehlerhafte Vorstellungen über die Entscheidungen der anderen Aktoren formal darzustellen: Solche fehlerhaften Vorstellungen können sich nämlich auf dem Hintergrund des Spielverlaufs verändern (vgl. etwa *Fudenberg / Kreps* 1992; *Fudenberg / Kreps* 1993; *Kalai / Lehrer* 1993; *Fudenberg / Levine* 1993). Damit wird Lernen in gewisser Weise als Anpassung von Realitätsdefinitionen modelliert, allerdings im Rahmen von formalen Regeln. Die Individuen haben keine Möglichkeit, die Regeln in ihren Kontexten zu reflektieren (vgl. *van Witteloostuijn* 1990). Damit gibt es auch keine höheren Reflexionsniveaus des individuellen Lernens im Sinne von *Bateson* (1992). Folglich entfällt auch die Grundlage für höhere Niveaus des organisationalen Lernens, nämlich Double-loop- und Deutero-Lernen.

5.3.2 Organisationales Lernen (Institutionalisierungsdimension)

Organisationales Lernen wurde definiert als institutionalisiertes individuelles Lernen, d.h. als die Etablierung systematischer kollektiver Verhaltensmuster. Im Unterschied zur Argumentation von Esser, Lindenberg (vgl. oben Abschnitt 4.3) und Frey (vgl. Abschnitt 4.4) kann man mit Hilfe der Spieltheorie zeigen, daß sich unter angebbaren Bedingungen kollektive Verhaltensmuster dauerhaft herausbilden. Diese werden nicht auf soziale Produktionsfunktionen als exogene Faktoren zurückgeführt bzw. funktionalistisch begründet. Sie entsprechen allerdings nur Institutionalisierungsmechanismen vom Typ Geltung 1.

Geltung 2 muß nach wie vor exogen eingeführt werden. Zum einen werden zwar „refinements“ wie etwa „focal points“, „Robustheit“, „pre-play communication“ oder „common knowledge“ als gemeinsame Situationsdefinitionen berücksichtigt. Sie können als Institutionen im Sinne von Geltung Typ 2a verstanden werden.

Darüber hinaus setzt die Tatsache, daß die Spieler ein Spiel spielen, die Akzeptanz von selbstverpflichtenden Spielregeln voraus. Dies betrifft die Geltung vom Typ 2b.[51] Die Einführung solcher Spielregeln ist im Rahmen der Spieltheorie

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

nicht mehr modellierbar (vgl. *Kliemt / Zimmerling* 1993).[52] Dazu müssen Modifikationen des homo oeconomicus, d.h. zusätzliche Eigenschaften der Spieler, eingeführt werden: Beispielsweise bezieht *Elster* (1989) die Wertrationalität als eine Eigenschaft der Spieler ein, und *Johnson* (1993) versucht, die selbstverpflichtende Wirkung von Kommunikation auf die Spieler aufzuzeigen. In beiden Fällen wird aber dadurch die formale Strenge des spieltheoretischen Modells aufgehoben.[53]

Die Leistung der Spieltheorie besteht zusammenfassend darin, das Institutionalisierungsproblem präzisiert zu haben: Sie zeigt auf, welche Modifikationen des Spiels oder des Spielers für ein rationales Individuum wünschenswert wären, ohne daß das rationale Individuum diese Modifikationen selbst erbringen kann (vgl. *Kliemt* 1991, S. 197). Zur Wissensdimension trägt die Spieltheorie allerdings wenig bei, weil sie von vornherein von einer strengen Rationalitätsdefinition ausgeht. Damit werden aus Gründen der Formalisierbarkeit wichtige empirisch relevante Probleme wie die Reflexion des eigenen Handelns oder die Entstehung von Innovationen ausgeschlossen.

6. Schlußfolgerungen

Was folgt aus den bisherigen Ausführungen für eine Theorie des organisationalen Lernens?

Zum *ersten:* Um über eine reifizierende Analogie zwischen individuellem und organisationalem Lernen hinauszukommen, sollte zwischen Wissensdimension und Institutionalisierungsdimension unterschieden werden. Die ökonomische Theorie hat bisher ausschließlich zur Erklärung der Institutionalisierungsdimension beigetragen, allerdings ist auch hier ihr Erklärungspotential auf die Institutionalisierung durch Sanktionierung bzw. durch Wahl von Rollensets beschränkt (Geltung 1 und 2a). Insbesondere die Spieltheorie zeigt, wie Institutionalisierung über dauerhafte Interaktion von Individuen erklärt werden kann. Die Grenzen der ökonomischen Theorie liegen darin, daß sie nicht erklären kann, wie Normen ohne die Existenz einer intrinsischen Motivation entstehen können (d.h. ohne Geltung 2b). Darüber hinaus wird die Wissensdimension nur als exogener Faktor diskutiert und damit die Innovationsforschung aus der Ökonomik ausgeschlossen. Die Managementlehre hat im Gegensatz dazu die Wissensdimension im Stufenkonzept des organisationalen Lernens ausführlich behandelt, jedoch die Institutionalisierungsdimension nur auf einer Ad-hoc-Ebene integriert.

---

51 Dies bedeutet, daß Institutionen im Sinn von Hobbes II ausgeschlossen bleiben (vgl. Fußnote 48).

52 Dies schließt nicht aus, daß die Akzeptanz von Spielregeln einen ökonomischen Wert darstellt.

53 Beide Autoren nehmen die Modifikation nicht im Sinne eines Sequenzmodells vor (vgl. oben Abschnitt 5.4), sondern versuchen, diese mit spieltheoretischen Argumenten zu verknüpfen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Zum *zweiten:* Eine theoretische Integration der Wissens- und der Institutionalisierungsdimensionen ist noch zu leisten. Eine Vorstufe dazu bilden die Ansätze, die das Verhältnis der beiden Dimensionen auf empirischer Ebene untersuchen (vgl. *Frey* 1992). Welche Konsequenzen auf der theoretischen Ebene dies haben könnte, macht die folgende Überlegung deutlich: Individuelles Lernen ist ein nicht umkehrbares Fortschreiten auf Ebenen zunehmender Reflexivität. Weil sich organisationales Lernen von individuellem Lernen durch Institutionalisierung unterscheidet, Institutionalisierung aber (wie in Abschnitt 3 gezeigt) eine Wahl zwischen den verschiedenen Geltungsmodi erfordert, ist organisationales Lernen nicht ein fortschreitender Prozeß. Vielmehr können — in Abhängigkeit von den trade-offs — verschiedene Geltungsmodi gewählt werden.[54]

Zum *dritten:* Obwohl eine theoretische Integration der Wissens- und Institutionalisierungsdimension noch aussteht, kann anhand einer aktuellen Kontroverse gezeigt werden, daß das organisationale Lernen tatsächlich auf verschiedenen Lernniveaus institutionalisiert werden kann (vgl. *Adler / Cole* 1993; *Berggren* 1994). Es geht dabei um die Gegenüberstellung von zwei prominenten Bespielen der Automobilproduktion, nämlich der zu Volvo gehörenden Fertigung in Uddevalla und des gemeinsamen Joint-ventures von Toyota und General Motors, der NUMMI-Fabrik in Kalifornien. In Uddevalla existierten teilautonome Arbeitsgruppen, in denen ganzheitliches und reflexives Lernen auf der Double-loop- und Deutero-Ebene möglich war.[55] Bei NUMMI hingegen gibt es kurzgetaktete Fließbänder, verknüpft mit Qualitätszirkeln. In diesen trainieren die Arbeiterinnen und Arbeiter, mit der Stoppuhr umzugehen, um innerhalb eines kleinen Bandabschnittes ihre eigene Arbeit zu rationalisieren. Dabei wird ohne Inanspruchnahme höherer Reflexionsniveaus und intrinsischer Motivation gelernt.

Adler / Cole (1993) stellen im Rahmen eines Vergleichs beider Konzepte die These auf, daß nur die standardisierte Arbeitsorganisation bei NUMMI organisationales Lernen ermögliche, wohingegen in teilautonomen Arbeitsgruppen bei Volvo lediglich das individuelle, aber eben nicht das organisationale Lernen gefördert würde. Die Standardisierung der Arbeitsprozesse sehen sie als unverzichtbare Voraussetzung dafür an, daß Wissen institutionalisiert wird. Da in den teilautonomen Arbeitsgruppen Standardisierungen und Dokumentationen der Arbeitsprozesse fehlten, könne das Wissen nicht diffundieren und damit nicht organisational gelernt werden (vgl. *Adler / Cole* 1993, S. 86, 92).

Hingegen argumentiert *Berggren* (1994), daß in teilautonomen Arbeitsgruppen die Möglichkeit besteht, höhere Lernniveaus zu erreichen. Systematische Personalrotationen und die Zusammenarbeit mit Ingenieuren und Managern erlauben die Diffusion von Wissen, das aus der Konfrontation mit anderen „frames" entstanden ist.

---

[54] Vgl. eine ähnliche Argumentation in bezug auf moralische Rückschritte einer Gesellschaft bei *Homann / Suchanek* (1992, S. 25ff.).

[55] Das Werk in Uddevalla ist 1993 geschlossen worden, jedoch besteht zwischen den Teilnehmern der Kontroverse Übereinstimmung darüber, daß dies nicht auf nachhaltige Wettbewerbsnachteile des Uddevalla-Konzeptes zurückzuführen ist.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Beide Autoren stimmen darin überein, daß die Lernprozesse bei Uddevalla zu mehr langfristiger Flexibilität führen, wohingegen die Erreichung kurzfristiger Effizienz und Qualitätsstandards mit dem NUMMI-System eher gewährleistet ist. Das Uddevalla-Konzept führt damit eher zu einer Kernkompetenz im Sinne eines schwer imitierbaren bzw. substituierbaren Wettbewerbsvorteils (vgl. Abschnitt 1 dieser Arbeit). Das NUMMI-Konzept ist dagegen eher für standardisierte Massenproduktion geeignet.[56]

Die Unterschiede zwischen den beiden Konzepten machen deutlich, daß innovatives Lernen allein aufgrund von extrinsischer Motivation nicht möglich ist: Wird nur aufgrund von Sanktionen gelernt, so wirkt dies nur, wenn für klar definierte Lernziele Belohnungen ausgesetzt werden. Innovationen, für die ein Lernziel nicht definiert werden kann, werden durch extrinsische Motivation nicht nur nicht gefördert, sondern unter bestimmten Umständen sogar verhindert. Es ist dies immer dann der Fall, wenn die extrinsische Motivation die intrinsische zerstört (vgl. Abschnitt 4.4 dieser Arbeit und *Osterloh / Gerhard* 1992).

Zum *vierten:* Unsere Überlegungen zum organisationalen Lernen können auch auf die Implementation ethischer Grundsätze bezogen werden: Die höheren Reflexionsniveaus beim Double-loop- und Deutero-Lernen implizieren nämlich immer auch eine Diskussion von Geltungsansprüchen: Wenn die Angemessenheit von „frames" neu bestimmt wird, steht gleichzeitg ihre normative Richtigkeit zur Disposition (vgl. *Habermas* 1983). Die Anforderungen an innovationsfreundliche Unternehmensstrukturen sind damit zugleich Anforderungen, die erfüllt sein müssen, um das Konzept der Unternehmensethik in Organisationen zu realisieren (vgl. *Steinmann / Gerhard* 1992). Dies hat seine Ursache darin, daß die Umsetzung ethischer Normen ebenso wie Produkt- und Prozeßinnovationen einen organisationalen Lernprozeß voraussetzt. Das heißt aber auch, daß organisationales Lernen immer mit der Diskussion über normative Geltung verknüpft ist.

Zum *fünften:* Die ökonomische Theorie und die Spieltheorie können das eingangs beschriebene Erscheinungsbild des organisationalen Lernens nur partiell in die eigene Sprache übersetzen. Deshalb ist es problematisch, wenn neuerdings die ökonomische Spieltheorie auf unmittelbar praktische organisatorische Fragestellungen angewandt wird (vgl. *Milgrom / Roberts* 1991; *McMillan* 1992). Prozesse des Double-loop- und Deutero-Lernens sowie der Institutionalisierung vom Typ Geltung 2b werden in den ökonomisch-spieltheoretischen Modellen nicht nur nicht thematisiert, sondern sie können auf dieser Basis auch nicht gestaltet werden.

Dies heißt für die Managementwissenschaft, daß sie die ökonomische Theorie als einen von mehreren möglichen Bezugsrahmen in ihre Analyse miteinbeziehen und mit anderen disziplinären Bezugsrahmen in ein „framework" mittlerer Abstraktionsweite einbauen sollte. Freilich können diese „frameworks" nicht die Eleganz des geschlossenen ökonomischen bzw. spieltheoretischen Weltbildes

[56] Vgl. auch das bei *Frei* (1993) geschilderte Beispiel.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

haben (vgl. *Osterloh / Grand* 1994). Dies ist der Preis, der zu zahlen ist, soll die Ökonomik und insbesondere die Spieltheorie nicht bloß „games economists play" (*Fisher* 1989) produzieren.

Literatur

Adler, Paul S. / *Cole* Robert E. (1993): Designed for Learning: A Tale of Two Auto Plants, in: Sloan Management Review, Jg. 35, Spring, S. 85-94.

Alchian, Armen A. / *Demsetz,* Harold (1972): Production, information costs, and economic organization, in: American Economic Review, 72, S. 777-795.

Apel, Karl-Otto (1992): Diskursethik vor der Problematik von Recht und Politik, in: Apel, Karl-Otto / Kettner, Matthias (Hrsg.): Zur Anwendung der Diskursethik in Politik, Recht und Wissenschaft, Frankfurt a.M., S. 29-61.

Argyris, Chris / *Schön,* Donald A. (1978): Organizational learning: A theory of action perspective, Reading, Mass. u.a.

Arrow, Kenneth J. (1991): Scale Returns in Communication and Elite Control of Organizations, in: Journal of Law and Economics and Organisation, Jg. 7, Special Issue, S. 1-16.

Axelrod, Robert (1991): Die Evolution der Kooperation, München.

Barney, Jay (1991): Firm Resources and Sustained Advantage, in: Journal of Management, Jg. 17, Nr. 1, S. 99-120.

Bartlett, Frederic C. (1932): Remembering: A study in experimental and social psychology, London.

Bateman, Thomas A. / *Zeithaml,* Carl P. (1989): The Psychological Context of Strategic Decisions: A Model and Convergent Experimental Findings, in: Strategic Management Journal, Jg. 10, S. 55-79.

Bateson, Gregory (1992): Ökologie des Geistes. Anthropologische, psychologische, biologische und epistemologische Perspektiven, 4. Aufl., Frankfurt a.M. (engl. Erstausgabe 1972).

Berggren, Christian (1994): Nummi vs. Uddevalla, in: Sloan Management Review, Jg. 36, Winter, S. 37-49.

Brennan, Geoffrey / *Buchanan,* James M. (1985): The Reason of Rules. Constitutional Political Economy, Cambridge.

Buchanan, James M. / *Vanberg,* Viktor (1991): The Market as a Creative Process, in: Economics and Philosophy, Jg. 7, S. 167-186.

Burns, James M. (1978): Leadership, New York.

Casson, Mark (1991): The Economics of Business Culture. Game Theory, Transanction Costs, and Economic Performance, Oxford.

Cohen, M. D. (1981): The Power of Parallel Thinking, in: Journal of Economic Behavior and Organization, Jg. 2, S. 285-306.

Cyert, Richard M. / *March,* James G. (1963): A Behavioral Theory of the Firm, Englewood Cliffs, NJ.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Davenport, Thomas H. (1993): Process Innovation, Reengineering Work through Information Technology, Boston Mass.

Davidow, William H. / *Malone,* Michael S. (1992): The Virtual Corporation. Structuring and Revitalizing the Corporation for the 21st Century, New York.

deGeus, Arie P. (1988): Planning as Learning, in: Harvard Business Review, Jg. 66, Nr. 2, S. 70-74.

Denzau, Arthur. T. / *North,* Douglass C. (1994): Shared mental-modells-ideologies and institutions, in: Kyklos, Jg. 47, Nr. 1, S. 3-31.

Diekmann, Andreas / *Preisendörfer,* Peter (1991): Umweltbewusstsein, ökonomische Anreize und Umweltverhalten, in: Schweizerische Zeitschrift für Soziologie, Jg. 17, Nr. 2, S. 207-231.

Dodgson, Mark (1993): Organizational learning: A review of some literatures, in: Organization Studies, Jg. 14, Nr. 3, S. 375-394.

Dörner, Dietrich (1979): Problemlösen als Informationsverarbeitung, 2. Aufl., Stuttgart.

Eichenberger, Reiner / *Frey,* Bruno S. (1993): ‚Superrationalität' oder: Vom rationalen Umgang mit dem Irrationalen, in: Jahrbuch für Neue Politische Ökonomie, Band 12, S. 50-84.

Elster, Jon (1989): The Cement of Society. A Study of Social Order, Cambridge.

Esser, Hartmut (1990): „Habits", „Frames", and „Rational Choice" — Die Reichweite von Theorien der rationalen Wahl (am Beispiel der Erklärung des Befragtenverhaltens), in: Zeitschrift für Soziologie ZfS, Jg. 19, Nr. 4, S. 231-247.

— (1991): Alltagshandeln und Verstehen. Zum Verhältnis von erklärender und verstehender Soziologie am Beispiel von Alfred Schütz und „Rational Choice", Tübingen.

Fiol, C. Marlene / *Lyles,* Marjorie A. (1985): Organizational Learning, in: Academy of Management Review, Jg. 10, Nr. 4, S. 803-813.

Fisher, Franklin M. (1989): Games economists play: A noncooperative view, in: RAND Journal of Economics, Jg. 20, Nr. 1, S. 113-124.

Föhr, Silvia / *Lenz,* Hansruedi (1992): Unternehmenskultur und ökonomische Theorie, in: Staehle, Wolfgang H. / Conrad, Peter (Hrsg.): Managementforschung 2, Berlin u.a. , S. 111-162.

Frank, Robert H. (1992): Melting Sociology and Economics: James Coleman's Foundation of Social Theory, in: Journal of Economic Literature, March, S. 147-170.

Frei, Felix (1993): Partizipation und Selbstregulation bei CIM: Das „Baugruppenprojekt" bei Alcatel STR, in: Cyranek, Günther / Ulich, Eberhard (Hrsg.): CIM — Herausforderung an Mensch, Technik, Organisation, Zürich, S. 321-338.

Frese, Erich (1993): Geschäftssegmentierung als organisatorisches Konzept — Zur Leitbildfunktion mittelständischer Strukturen für Großunternehmungen, in: Zeitschrift für betriebswirtschaftliche Forschung, Jg. 45, S. 999-1024.

Frey, Bruno S. (1990): Ökonomie ist Sozialwissenschaft. Die Anwendung der Ökonomie auf neue Gebiete, München.

— (1992): Tertium Datur: Pricing Regulating and Intrinsic Motivation, Kyklos Jg. 45, S. 161-184.

— (1994): How Intrinsic Motivation is Crowded Out and In, in: Rationality and Society, im Druck.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Frey, Bruno S. / *Eichenberger,* Rainer (1989): Zur Bedeutung entscheidungstheoretischer Anomalien für die Ökonomik, in: Jahrbuch für Nationalökonomie und Statistik Bd. 206/2, Stuttgart.

Fudenberg, Drew / *Kreps,* David M. (1992): Learning Mixed Equilibrium, mimeo MIT.

—/— (1993): Learning in Extensive-Form Games, mimeo Harvard University.

Fudenberg, Drew / *Levine,* David K. (1993): Steady State Learning and Nash Equilibrium, in: Econometrica, Jg. 61, Nr. 3, S. 547-573.

Fudenberg, Drew / *Tirole,* Jean (1987): Non-cooperative Game Theory for Industrial Organization: An Introduction and Overview, in: Schmalensee, Richard / Willig, Richard D. (Hrsg.) (1989): Handbook of Industrial Organization, Boston.

—/— (1991): Game Theory, Cambridge Mass.

Gäfgen, Gérard (1983): Institutioneller Wandel und ökonomische Erklärung, in: Jahrbuch für Neue Politische Ökonomie, Band 2, S. 19-49.

Giddens, Anthony (1988): Die Konstitution der Gesellschaft. Grundzüge einer Theorie der Strukturierung. Mit einer Einführung von Hans Joas, Frankfurt a.M.

Habermas, Jürgen (1981): Theorie des kommunikativen Handelns, Bd. I u. II, Frankfurt a.M.

— (1983): Moralbewußtsein und kommunikatives Handeln, Frankfurt a.M.

— (1992): Faktizität und Geltung, Frankfurt a.M.

Haferkamp, Hans (Hrsg.) (1990): Sozialstruktur und Kultur, Frankfurt a.M.

Hammer, Michael / *Champy,* James (1993): Reengineering the Corporation, New York.

Hedberg, Bo (1981): How organizations learn and unlearn, in: Nystrom, Paul C. / Starbuck, William H. (Hrsg.): Handbook of Organizational Design, Oxford, Jg. 1, S. 3-27.

Hofer, Charles W. / *Schendel,* Dan (1978): Strategy Formulation: Analytical Concepts, St. Paul.

Homann, Karl (1988): Philosophie und Ökonomik. Bemerkungen zur Interdisziplinarität, in: Jahrbuch für Neue Politische Ökonomie, Band 7, S. 99-127.

Homann, Karl / *Suchanek,* Andreas (1989): Methodologische Überlegungen zum ökonomischen Imperialismus, in: Analyse & Kritik, II, S. 70-93.

—/— (1992): Grenzen der Anwendbarkeit einer „Logik des kollektiven Handelns", in: Schubert, Klaus (Hrsg.): Leistungen und Grenzen politisch-ökonomischer Theorie. Eine kritische Betrachtung zu Mancur Olson, Darmstadt, S. 13-27.

Homann, Karl / *Pies,* Ingo (1994): Wirtschaftsethik in der Moderne: Zur ökonomischen Theorie der Moral, in: Ethik und Sozialwissenschaften (EUS). Streitforum für Erwägungskultur, Jg. 5, Nr. 1, S. 3-12.

Huber, George P. (1991): Organizational Learning: The Contributing Processes and the Literatures, in: Organization Science, Jg. 2, Nr. 1, S. 88-115.

Isaacs, William N. (1993): Talking Flight: Dialogue, Collective Thinking, and Organizational Learning, in: Organizational Dynamics, Jg. 22, Nr. 2, S. 24-39.

Janis, Irving L. (1982): Groupthink, 2. Aufl., Boston.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Johnson, James (1993): Is Talk really cheap? Prompting Conversation between Critical Theory and Rational Choice, in: American Political Science Review, Jg. 87, Nr. 1, S.74-86.

Kahneman, Daniel / *Tversky,* Amos (1984): Choice, Values, and Frames, in: American Psychologist, Jg. 39, Nr. 4, S. 342-350.

Kalai, Ehud / *Lehrer,* Ehud (1993): Rational Learning leads to Nash Equilibrium, in: Econometrica, Jg. 61, Nr. 5, S. 1019-1045.

Keller, Monika (1976): Kognitive Entwicklung und soziale Kompetenz. Zur Entstehung der Rollenübernahme in der Familie und ihre Bedeutung für den Schulerfolg, Stuttgart.

Kets de Vries, Manfred F. R. / *Miller,* Danny (1984): The Neurotic Organization: Diagnosing and Changing Counterproductive Styles of Management, San Francisco.

Kirchgässner, Gebhard (1991): Homo oeconomicus. Das ökonomische Modell individuellen Verhaltens und seine Anwendung in den Wirtschafts- und Sozialwissenschaften, Tübingen.

Kliemt, Hartmut, (1991): Der Homo oeconomicus in der Klemme, in: Esser, Hartmut / Troitzsch, Klaus G. (Hrsg.): Modellierung sozialer Prozesse, Sozialwissenschaftliche Tagungsberichte, Informationszentrum Sozialwissenschaften, Bonn, S. 179-204.

Kliemt, Hartmut / *Zimmerling,* Ruth (1993): Quo vadis Homo Oeconomicus? Über einige neuere Versuche, das Modell eines Homo Oeconomicus fortzuentwickeln, in: Homo Oeconomicus I, München, S. 1-37.

Knight, Frank H. (1921): Risk, Uncertainty and Profit, Neuauflage 1971, Chicago.

Kreps, David (1990): Corporate Culture and Economic Theory, in: Alt, J. E. / Shepsle, K. A. (Hrsg.): Percpectives on positive political economy, Cambridge, S. 90-143.

— (1991): A Course in Microeconomic Theory, Princeton Levitt.

Levitt, Barbara / *March,* James G. (1988): Organizational Learning, in: Annual Review of Sociology, Jg. 10, Nr. 4, S. 319-340.

Lindenberg, Siegwart (1990): Rationalität und Kultur. Die verhaltenstheoretische Basis des Einflusses von Kultur auf Transaktionen, in: Haferkamp, Hans (Hrsg.): Sozialstruktur und Kultur, Frankfurt a.M., S. 249-287.

Löhr, Albert (1991): Unternehmensethik und Betriebswirtschaftslehre, Stuttgart.

Löhr, Albert / *Osterloh,* Margit (1993): Ökonomik und Ethik als Grundlage organisationaler Beziehungen, in: Staehle, Wolfgang H. / Sydow, Jörg (Hrsg.): Managementforschung 3, Berlin u.a., S. 109 - 155.

March, James G. / *Sproull,* Lee / *Tamuz,* Michal (1991): Learning from samples of one or fewer, in: Organization Science, Jg. 2, Nr. 1, S. 1-13.

Maturana, Humberto R. / *Varela,* Francisco (1984): Der Baum der Erkenntnis, Bern u.a.

Mayer, Thomas (1993): Truth versus Precision in Economics, Cambridge.

McCaskey, Michael B. (1991): Mapping: creating, maintaining, and relinquishing conceptual frameworks, in: Henry, Jane (Hrsg.): Creative Management, London u.a., S. 135-152.

McMillan, John (1992): Games, Strategies, and Managers, New York u.a.

Meyer, John W. / *Rowan,* Brian (1977): Institutionalized Organizations: Formal Structure as Myth and Ceremony, in: American Journal of Sociology, Jg. 83, Nr. 2, S. 340-363.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Milgrom, Paul / *Roberts,* John (1991): Economics, Organization, and Management, Englewood Cliffs, New Jersey.

Miller, Georg A. / *Galanter,* Eugene / *Pribram,* Karl H. (1960): Plans and the Structure of Behavior, New York.

Minkler, Alanson P. (1993): The Problem with Dispersed Knowledge: Firms in Theory and Practice, in: Kyklos, Jg. 46, Nr. 4, S. 569-587.

Minsky, Marvin, L. (1968): Semantic Processing, Cambridge Mass.

Moorhead, Gregory / *Ference,* Richard / *Neck,* Chris P. (1991): Group Decision Fiascoes Continue: Space Shuttle Challenger and a Revised Groupthink Framework, in: Human Relation, Jg. 44, Nr. 6, S. 539-550.

Neisser, Ulric (1979): Kognition und Wirklichkeit. Prinzipien und Implikationen der kognitiven Psychologie, Stuttgart.

Nelson, Richard R. (1991): Why Do Firms Differ, and How Does It Matter?, in: Strategic Management Journal, Jg. 12, S. 61-74.

Nelson, Richard R. / *Winter,* Sidney G. (1982): An Evolutionary Theory of Economic Change, Cambridge, Mass.

Nonaka, Ikujiro (1992): Wie japanische Konzerne Wissen erzeugen, in: Harvardmanager, Nr. 2, S. 95-103.

North, Douglass C. (1988): Theorie des institutionellen Wandels. Eine neue Sicht der Wirtschaftsgeschichte, Tübingen.

— (1992): Institutionen, institutioneller Wandel und Wirtschaftsleistung, Tübingen.

Nystrom, Paul C. / *Starbuck,* William H. (1984): To Avoid Organizational Crises, Unlearn, in: Organizational Dynamics, Jg. 13, Nr. 4, S. 53-65.

Olson, Mancur (1968): Die Logik des kollektiven Handelns. Kollektivgüter und die Theorie der Gruppen, Tübingen.

Opp, Karl-Dieter (1983): Die Entstehung sozialer Normen, Tübingen.

Ordelheide, Dieter / *Rudolph,* Bernd / *Büsselmann,* Elke (Hrsg.) (1991): Betriebswirtschaftslehre und ökonomische Theorie, Stuttgart.

Osterloh, Margit (1991): Managementrolle: Alleinentscheider, in: Staehle, Wolfgang H. (Hrsg.): Handbuch Management. Die 24 Rollen der exzellenten Führungskraft, Wiesbaden, S. 243-261.

— (1994): Kulturalismus versus Universalismus. Reflektionen zu einem Grundlagenproblem des interkulturellen Managements, in: Schiemenz, Bernd / Wurl, Hans Jürgen (Hrsg.): Internationales Management. Beiträge zur Zusammenarbeit. Festschrift für Eberhard Dülfer zum 70. Geburtstag, Wiesbaden, S. 95-116.

Osterloh, Margit / *Frost,* Jetta (1994): Business Reengineering — Modeerscheinung oder „Business Revolution"?, in: Zeitschrift Führung + Organisation (zfo), erscheint demnächst.

Osterloh, Margit / *Gerhard,* Brigit (1992): Neue Technologien, Arbeitsanforderungen und Aufgabenorientierung: Zum Verhältnis von intrinsischer und extrinsischer Motivation, in: Lattmann, Charles / Probst, Gilbert J.B. / Tapernoux, Frédéric (Hrsg.) (1992): Die Förderung der Leistungsbereitschaft des Mitarbeiters als Aufgabe der Unternehmensführung, Heidelberg, S. 117-134.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Osterloh, Margit / *Grand,* Simon (1994): Modellbildung versus Frameworking: Zu den Ebenen betriebswirtschaftlicher Theoriebildung, Manuskript Institut für betriebswirtschaftliche Forschung, Zürich.

Pautzke, Gunnar (1989): Die Evolution der organisatorischen Wissensbasis. Bausteine zu einer Theorie des organisatorischen Lernens, München.

Piaget, Jean (1973): Einführung in die genetische Erkenntnistheorie, Frankfurt a.M.

Polanyi, Michael (1967): The tacit dimension, Garden City (New York).

Porter, Michael E. (1991): Towards a Dynamic Theory of Strategy, in: Strategic Management Journal, Jg. 12, Special Issue, Winter, S. 95-117.

Prahalad, Coimbatore K. / *Hamel,* Gary (1991): Nur Kernkompetenzen sichern das Überleben, in: Harvardmanager, Nr. 2, S. 66-78.

Presidential Commission (1986): Report of the Presidential Commission on the Space Shuttle Accident, 5 Vol., Washington D.C.

Rosenberg, Nathan (1976): Perspectives on Technology, Cambridge.

Rubinstein, Ariel (1991): Comments on the Interpretation of Game Theory, in: Econometrica, Jg. 59, Nr. 4, S. 909-924.

Rumelhart, Norman (1980): Schemata: The Building-Blocks of Cognition, in: Spiro, R. / Bruce, B. / Brewer, W. (Hrsg.): Theoretical Issues in Reading Comprehension, Hillsdale, N.J.

Sandler, Todd (1992): Collective Action. Theory Applications, Ann Arbor.

Sattelberger, Thomas (Hrsg.) (1991): Die lernende Organisation. Konzepte für eine neue Qualität der Unternehmensentwicklung, Wiesbaden.

Schank, Roger / *Abelson,* Robert (1977): Scripts, Plans, Goals, and Understanding: An inquiry into human knowledge structures, Hillsdale, N.J.

Schein, Edgar H. (1986): Organizational Culture and Leadership, San Francisco u.a.

— (1993): On Dialogue, Culture, and Organizational Learning, in: Organizational Dynamics, Jg. 22, Nr. 2, S. 40-51.

Schelling, Thomas (1960): The Strategy of Conflict, Boston.

Schmookler, Jacob (1966): Invention and Economic Growth, Cambridge Massachusetts.

Schneider, Susan C. / *Angelmar,* Reinhard (1993): Cognition in Organizational Ananlysis: Who's Minding the Store?, in: Organization Studies, Jg. 14, Nr.3, S. 347-374.

Schüßler, Rudolf (1990): Kooperation unter Egoisten: Vier Dilemmata, München.

Schwarz, Norbert (1985): Theorien konzeptgesteuerter Informationsverarbeitung in der Sozialpsychologie, in: Frey, Dieter / Irle, Martin (Hrsg.): Theorien der Sozialpsychologie, Band III: Motivations- und Informationsverarbeitungstheorien, Bern u.a., S. 269-291.

Sen, Amartya (1987): On Ethics and Economics, Oxford.

Senge, Peter M. (1990): The fifth discipline. The art and practice of the learning organization, London u.a.

— (1990a): The Leader's New Work: Building Learning Organizations, in: Sloan Management Review, Jg. 32, S. 7-23.

— (1993): Communities of Commitment: The Heart of Learning Organizations, in: Organizational Dynamics, Jg. 22, Nr. 2, S. 5-23.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Shrivastava, Paul (1983): A Typology of Organizational Learning Systems, in: Journal of Management Studies, Jg. 20, Nr. 1, S. 7-28.

Siegenthaler, Hansjörg (1993): Regelvertrauen, Prosperität und Krisen, Tübingen.

Simon, Herbert A. (1951): A Formal Theory of the Employment Relationship, in: Econometrica, Jg. 19, S. 293-305.

— (1961): Administrative Behavior, 2nd ed., New York.

— (1991): Bounded rationality and organizational learning, in: Organization Science, Jg. 2, Nr. 1, S. 125-134.

Sims, Henry P. jr. / *Gioia,* Dennis A. (1986): The Thinking Organization, San Francisco.

Srubar, Ilja (1992): Grenzen des „Rational-Choice"-Ansatzes, in: Zeitschrift für Soziologie ZfS, Jg. 21, Nr. 3, S. 157-165.

Starbuck, William H. (1984): Organizations as Action Generators, in: American Sociological Review, Jg. 48, S. 91-102.

Stata, Ray (1989): Organizational Learning — The Key to Management Inovation, in: Sloan Management Review, Jg. 31, Nr. 3, S. 63-74.

Steinmann, Horst / *Gerhard,* Birgit (1992): Effizienz und Ethik in der Unternehmensführung, in: Homann, Karl (Hrsg.): Aktuelle Probleme der Wirtschaftsethik, Berlin, S. 159-182.

Steinmann, Horst / *Hennemann,* Carola (1993): Personalmanagementlehre zwischen Managementpraxis und mikro-ökonomischer Theorie, Diskussionsbeitrag Nr. 76, Lehrstuhl für Allg. Betriebswirtschaftslehre und Unternehmensführung der Universität Erlangen-Nürnberg.

Steinmann, Horst / *Schreyögg,* Georg (1986): Zur organisatorischen Umsetzung der Strategischen Kontrolle, in: Zeitschrift für betriebswirtschaftliche Forschung, Jg. 38, Nr. 9, S. 747-765.

Suchanek, Andreas (1991): Der ökonomische Ansatz und das Verhältnis von Mensch, Institution und Erkenntnis, in: Bievert, Bernd / Held, Martin (Hrsg.): Das Menschenbild der ökonomischen Theorie, Frankfurt a.M. u.a., S. 76-93.

Suchman, Lucy A.(1987): Plans and situated actions. The problem of human machine-communications, Cambridge.

Sugden, Robert (1989): The Economic Theory of Rights, Cooperation and Welfare, Blackwell.

Taylor, Michael (1976): Anarchy and Cooperation, London.

Teece, David J. / *Pisano,* Gary / *Shuen,* Any (1992): Dynamic Capabilities and Strategic Management, Working Paper, Harvard mimeo.

Tolman, Edward (1948): Cognitive maps in rats and man, in: Psychological Review, Jg. 55, S. 189-208.

Tversky, Amos / *Kahneman,* Daniel (1981): The Framing of Decisions and the Psychology of Choice, in: Science, Jg. 221, S. 453-458.

—/— (1984): siehe unter: Kahneman, Daniel / Tversky, Amos (1984).

Ulrich, Hans (1988): Von der Betriebswirtschaftslehre zur systemorientierten Managementlehre, in: Wunderer, Rolf (Hrsg.): Betriebswirtschaftslehre als Management- und Führungslehre, 2. Aufl., Stuttgart, S. 173-190.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Vanberg, Viktor (1983): Der individualistische Ansatz zu einer Theorie der Entstehung und Entwicklung von Institutionen, in: Jahrbuch für Neue Politische Oekonomie, Band 2, Tübingen. S. 50-69.

— (1984): „Unsichtbare-Hand Erklärung" und soziale Normen, in: Todt, Horst (Hrsg.): Normengeleitetes Verhalten in den Sozialwissenschaften, Berlin, S. 115-148.

— (1989): Rational choice, Role Following, and Institutions: An Evolutionary Perspective, Mimeo.

van Witteloostuijn, Arjen (1990): Learning in Economic Theory: A Taxonomy with an Application to Expectations Formation, in: Journal of Economic Psychology, Jg. 11, S. 183-207.

Weick, Karl E. (1969): The social psychology of organizing, Reading Mass.

— (1991): The nontraditional quality of organizational learning, in: Organization Science, Jg. 2, Nr. 1, S. 116-124.

Winograd, Terry / *Flores,* Fernando (1992): Erkenntnis Maschinen Verstehen. Zur Neugestaltung von Computersystemen, 2. Aufl., Berlin.

Womak, James P. / *Jones,* Daniel T. / *Roos,* Daniel (1990): The Machine that Changed the World, New York.

zu Knyphausen, Dodo (1994): „Why are Firms different"? Der „Ressourcenorientierte Ansatz" im Mittelpunkt einer aktuellen Kontroverse im Strategischen Management, in: Die Betriebswirtschaft, Jg. 53, Nr. 6, S. 771-792.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Zur Anwendung wissenschaftstheoretischer Methoden auf ethische Theorien: Überlegungen am Beispiel des klassischen Utilitarismus

Von *Ulrich Gähde,* Bayreuth

1. Die Problemstellung

In der deutschen akademischen Philosophie hat sich eine relativ deutliche Aufgabenteilung zwischen dem theoretischen und dem praktischen Zweig entwikkelt: Ethik, politische Philosophie und Rechtsphilosophie einerseits, Logik, Wissenschaftstheorie, Erkenntnistheorie und Ontologie andererseits werden von zwei klar getrennten Gruppen von Wissenschaftlern bearbeitet. Der Informationsaustausch innerhalb dieser beiden nahezu disjunkten Gruppen ist vergleichsweise intensiv, der Austausch über die Grenze zwischen der theoretischen und der praktischen Philosophie hinweg dagegen marginal.

Die Trennung von theoretischer und praktischer Philosophie hat insbesondere dazu geführt, daß Methoden und Erkenntnisse der modernen Wissenschaftstheorie in der Ethik kaum rezipiert wurden. Dadurch sind wesentliche Parallelen zwischen dem Aufbau empirisch-deskriptiver Theorien einerseits und ethisch-normativer Theorien andererseits bisher unanalysiert geblieben. Erst in den letzten Jahren entwickelte sich die Untersuchung struktureller Ähnlichkeiten zwischen beiden Theoriearten langsam zu einem eigenen philosophischen Forschungsbereich. Beispiele für einschlägige Forschungsthemen sind etwa 1. die vergleichende Analyse von Rationalitätskonzeptionen in empirischen und ethischen Theorien (vgl. *Hollis / Vossenkuhl* [1992]; *Spohn* [1993]), 2. die wissenschaftstheoretische Untersuchung der Rolle von Idealisierungen und Approximationen bei der Computersimulation kooperativen Verhaltens mit Hilfe zellulärer Automaten (vgl. *Hegselmann* [1994]) sowie 3. die Problemstellung, um die es in diesem Vortrag geht: die vergleichende Analyse von Verteidigungsmechanismen bei deskriptiven und präskriptiven Theorien.

Grundgedanke ist dabei, daß sich ethische Theorien gegen Konflikte mit basalen moralischen Intuitionen in einer in manchen Hinsichten ähnlichen Weise verteidigen wie empirisch-deskriptive Theorien gegen konfligierende Beobachtungsdaten. Diese These wird am Beispiel ethischer Theorien illustriert, die zum Paradigma des klassischen Utilitarismus gehören.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Im folgenden Abschnitt 2 wird zunächst — aus wissenschaftstheoretischer Sicht — eine kurze Zusammenfassung von Verteidigungsmechanismen deskriptiver Theorien gegeben. Anschließend wird gezeigt, wie diese Überlegungen auf den Fall klassisch-utilitaristischer Theorien übertragen werden können (Abschnitte 3-5). Insbesondere wird auf den Zusammenhang zwischen den Systematisierungsleistungen ethischer Theorien hinsichtlich grundlegender Moralurteile und der Entwicklung ihres begrifflichen Apparats eingegangen. Dieser Zusammenhang wird am Beispiel der mehrfachen Bedeutungsverschiebungen illustriert, die der *Individuen-* sowie der *Nutzenbegriff* in der klassischen Entwicklungsphase des Utilitarismus erfahren haben (Abschnitte 6.1 bzw. 6.2). Weiterhin wird die Rolle behandelt, die empirische Vortheorien bei den Verteidigungsanstrengungen utilitaristischer Theorien gespielt haben (Abschnitt 7). In Abschnitt 8 wird ein Ausblick darauf gegeben, wie unter Verwendung des neueren wissenschaftstheoretischen Konzepts sog. *Theorien-Netze* die zuvor geschilderten qualitativen Überlegungen auf die mathematisierten Nachfolgetheorien des Utilitarismus im zwanzigsten Jahrhundert übertragen werden können. Den Abschluß bildet eine kurze Bemerkung zu den metaethischen Voraussetzungen dieser Überlegungen (Abschnitt 9).

2. Verteidigungsmechanismen bei empirisch-deskriptiven Theorien — Eine Zusammenfassung

Empirische Theorien sind Konflikten mit der Erfahrungsbasis keineswegs wehrlos ausgeliefert: Sie verfügen über eine Vielzahl eingebauter Verteidigungsmechanismen, mit denen sie sich gegen Kollisionen mit Beobachtungs- bzw. Meßdaten zur Wehr setzen können. Diese Mechanismen ermöglichen es ihnen inbesondere, ihre Fundamentalprinzipien in erheblichem Maße auf Kosten spezieller Gesetzmäßigkeiten zu immunisieren.

Auf diesen Sachverhalt ist bereits 1906 in einem Klassiker der wissenschaftstheoretischen Literatur, in Pierre Duhems *Ziel und Struktur einer physikalischen Theorie,* nachdrücklich hingewiesen worden (*Duhem* [1906, dt. 1908]). Duhem hob insbesondere einen Sachverhalt hervor, der wie folgt zusammengefaßt werden kann: Aus dem Scheitern einer Hypothese C, die zuvor aus einer Gesetzesmenge A sowie einer Menge B von Hilfshypothesen und Randbedingungen abgeleitet wurde, kann nur auf die Ungültigkeit der Konjunktion $A \wedge C$, nicht jedoch auf die Ungültigkeit eines der beiden Konjunktionsglieder geschlossen werden.

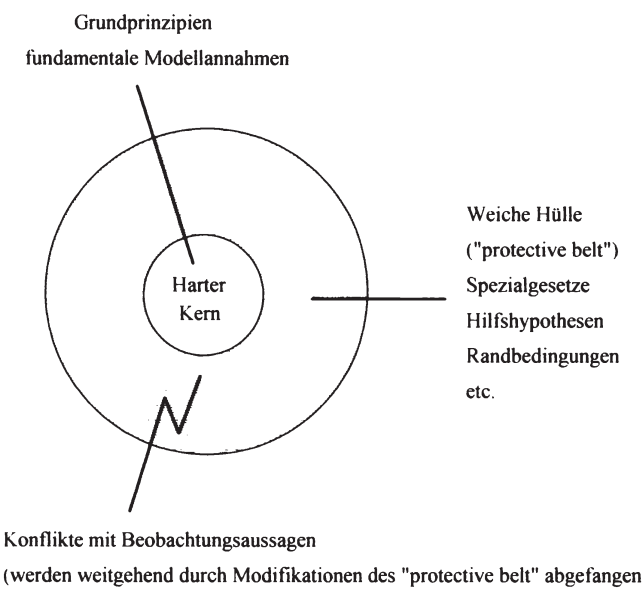
$$\frac{\begin{array}{l}(A \wedge B) \rightarrow C \\ \neg C\end{array}}{\neg (A \wedge B)}$$

Dadurch entsteht die Möglichkeit, zumindest versuchsweise die in A enthaltenen Fundamentalgesetze von einer Revision auszunehmen bzw. gegen Modifikatio-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

nen zu immunisieren. Diese Verteidigungsmöglichkeit, über die auch hochentwickelte empirische Theorien verfügen, steht offenbar in direktem Gegensatz zu Forderungen einer falsifikationistischen wissenschaftstheoretischen Position, wie sie etwa 30 Jahre später Karl Popper in *Logik der Forschung* vertreten hat.[1]

Schon Poppers Nachfolger Imre Lakatos zeichnet im Rahmen seiner *Methodologie der Forschungsprogramme* ein wesentlich andersartiges Bild vom Aufbau empirischer Theorien, das insbesondere erheblich erweiterte Verteidigungsmöglichkeiten für diese Theorien vorsieht (*Lakatos / Musgrave* [1974]). *Forschungsprogramme* im Lakatosschen Sinn sind als Theorienfolgen zu interpretieren. Die einzelnen Glieder dieser Folge weisen einen gemeinsamen *harten Kern* auf. Dieser besteht aus Grundgesetzen und fundamentalen Modellannahmen. Um den harten Kern legt sich eine *weiche Hülle,* die speziellere gesetzesartige Aussagen und Hilfshypothesen enthält. Diese fungiert als Schutzgürtel („protective belt"): Auf den Kern gerichtete Angriffe (in Form von Konflikten mit Beobachtungs- bzw. Meßdaten) können durch Umgestaltung der spezielleren Theorieteile weitgehend abgefangen werden. Die *weiche Hülle* wird dabei ständig modifiziert; sie



[1] Popper hat diesem Einwand u.a. dadurch Rechnung zu tragen versucht, daß er zwischen *zulässigen* und *unzulässigen* Hilfshypothesen unterschieden hat: Zulässig sind für ihn nur solche Hilfshypothesen, die den *Falsifizierbarkeitsgrad* einer Theorie erhöhen. Der (komparative) Begriff *Falsifizierbarkeitsgrad* sollte dabei unter Verwendung von *Inklusionsbeziehungen* zwischen Klassen potentiell falsifizierender Basissätze sowie *Dimensionsvergleichen* zwischen Hypothesen expliziert werden; vgl. *Popper* ([8]1984, insbesondere Kapitel VI).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

variiert im Rahmen eines Forschungsprogramms von Theorie zu Theorie. Die Regeln, die ihren Umbau steuern, werden von Lakatos in einer *positiven* sowie einer *negativen Heuristik* zusammengefaßt. Die positive Heuristik fordert, den Kern solange wie möglich gegen Widerlegungsversuche zu immunisieren. Die Regeln der negativen Heuristik beschreiben, in welcher Weise dies durch Umgestaltung des *protective belt* geschehen kann.

Ein weitgehend ähnliches Bild liegt auch zahlreichen modernen Konzeptionen der formalen Wissenschaftstheorie zugrunde. Ein Beispiel hierfür ist die sog. *strukturalistische Konzeption,* eine Variante der semantischen Theorienauffassung, die in vielen Hinsichten als eine Weiterentwicklung und formale Präzisierung Lakatosscher Ideen interpretiert werden kann (vgl. etwa *Stegmüller* [1986]). Im Rahmen dieses Ansatzes wurde der Versuch unternommen, das bei Lakatos nur vergleichsweise vage skizzierte Zusammenspiel zwischen dem *hartem Kern* einer Theorie und den zur *weichen Hülle* gehörenden peripheren Theorieteilen für eine Vielzahl konkreter empirischer Theorien im Detail zu analysien und mit Hilfe formaler metatheoretischer Konzepte zu erfassen. Bei diesen Analysen wurde deutlich, daß bei der Kooperation von fundamentalen und spezialisierten Theorieteilen sog. *theoretische Terme* eine Schlüsselrolle spielen. Diese Begriffe, die in der Wissenschaftstheorie bereits seit mehr als fünfzig Jahren diskutiert werden (vgl. etwa *Carnap* [1966, dt. 1986, S. 225 ff.]), sind u.a. durch die folgenden beiden formalen Bedingungen charakterisiert: Erstens treten sie in den Fundamentalprinzipien einer Theorie wesentlich auf, bleiben dort jeoch in ihrer Bedeutung weitgehend unterbestimmt. Zweitens wird eine nähere Bedeutungszuweisung erst durch das Zusammenspiel von Fundamentalprinzipien mit ergänzenden Spezialgesetzen, Hilfshypothesen, Randbedingungen etc. bewirkt (vgl. *Gähde* [1990]). Diese spezielleren Aussagen sind — in Lakatos' Terminologie — zu den peripheren, den *protective belt* konstituierenden Theorieteilen zu rechnen.

Die Rolle, die theoretische Terme bei den Verteidigungsanstrengungen empirisch-deskriptiver Theorien spielen, kann wie folgt zusammengefaßt werden: Angenommen, es kommt zu einem Konflikt zwischen *Theorie und Erfahrung;* eine mit Hilfe einer empirischen Theorie formulierte Hypothese scheitert an der verfügbaren Datenbasis. Dann wird man zunächst versuchen, den Konflikt dadurch auszuräumen, daß man Teile der Theoriehülle modifiziert, die fundamentalen Gesetzmäßigkeiten dagegen unangetastet läßt. Durch die Modifikation speziellerer, zur *weichen Hülle* gehörender Theorieteile ändert sich im allgemeinen jedoch zugleich die Bedeutung der theoretischen Terme. Dies wiederum kann eine Bedeutungsänderung der Fundamentalgesetze (d.h. eine Änderung der Wahrheitsbedingungen dieser Gesetze) induzieren — ohne daß diese dazu ihre äußere Gestalt auch nur im mindesten zu ändern bräuchten.

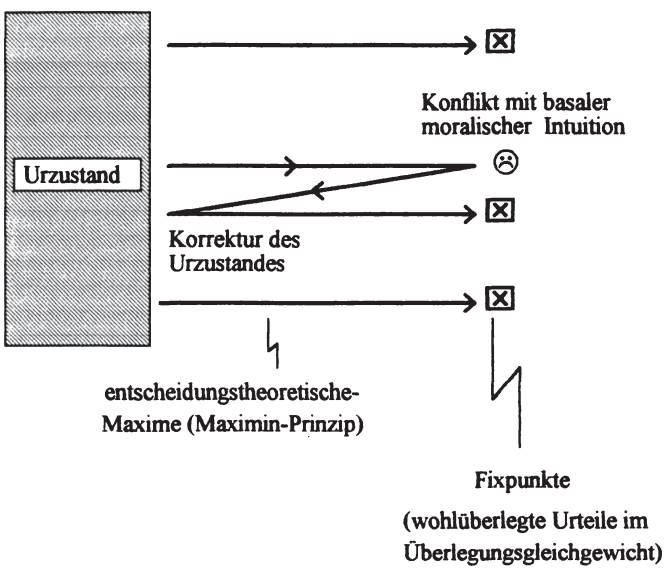
Diese Mechanismen sind in den letzten Jahren für eine Vielzahl insbesondere naturwissenschaftlicher Theorien im Detail untersucht worden. Ein Beispiel für eine derartige Analyse findet sich in *Gähde / Stegmüller* (1986).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

3. Grundidee der Übertragung auf ethische Theorien

Im folgenden wird die These aufgestellt, daß bei der Verteidigung ethischer Theorien gegen Konflikte mit basalen moralischen Intuitionen weitgehend analoge Mechanismen wirken.

Dabei wird davon ausgegangen, daß in den meisten ethischen Konzeptionen grundlegenden moralischen Urteilen explizit oder implizit eine gewisse Korrektivkraft zugesprochen wird:[2] Konflikte zwischen den normativen Implikationen ethischer Theorien und basalen Moralurteilen werden in aller Regel nicht einfach hingenommen. Statt dessen werden sie dazu führen, daß sowohl die betreffende Theorie als auch die konfligierenden moralischen Intuitionen überprüft werden: Moralische Intuitionen können dabei nicht nur die Änderung einer ethischen Theorie mit konfligierenden normativen Implikationen erzwingen, sondern umgekehrt ihrerseits im Licht einer ethischen Theorie ihre ursprüngliche Überzeugungskraft verlieren. Bekanntlich hat John Rawls vorgeschlagen, dieses Wechselspiel mit der Konzeption eines *reflective equilibrium* zu erfassen. Zugleich hat er versucht, basale moralische Intuitionen zu Fragen der (distributiven) Gerechtigkeit im Rahmen einer vertragstheoretischen Konzeption in einen systematischen Zusammenhang zu bringen (vgl. *Rawls* [1971, dt. 1975, S. 38 sowie S. 68-71]).



[2] Zu den metaethischen Voraussetzungen der folgenden Überlegungen vgl. Abschnitt 9.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Mit der im folgenden vorgenommenen Parallelisierung deskriptiver und normativer Theorien hinsichtlich ihrer Systematisierungsleistung sollen die fundamentalen Unterschiede zwischen beiden Theorietypen keineswegs überspielt werden. Insbesondere sind auch basale moralische Überzeugungen offensichtlich sowohl (in zeitlicher Hinsicht) innerpersonell wesentlich instabiler als auch interpersonell weniger konsensfähig als Überzeugungen über Beobachtbares: Sie bilden keine Datenbasis, die sich hinsichtlich ihrer Festigkeit und Kohärenz mit der Beobachtungsbasis empirischer Theorien vergleichen könnte.[3]

Angenommen jedoch, ein bestimmtes intuitives Moralurteil behält auch angesichts der Konfrontation mit einer ethischen Theorie seine anfängliche Plausibilität und interpersonelle Konsensfähigkeit bei: Der Konflikt wird zugunsten der moralischen Intuition entschieden. In diesem Fall wird man (analog wie im Fall empirisch-deskriptiver Theorien) bemüht sein, die ethische Theorie in einer Weise zu modifizieren, die ihre Fundamentalprinzipien unberührt läßt. Man wird statt dessen versuchen, eine Korrektur durch Modifikation peripherer Theorieteile zu erreichen, die in Lakatos' Terminologie zum *protective belt* der betreffenden Konzeption gehören. Dabei werden im allgemeinen die Grundterme der Theorie eine Folge von Bedeutungsverschiebungen durchlaufen, die ihrerseits Bedeutungsverschiebungen der äußerlich unveränderten Grundprinzipien induzieren können.

Im folgenden wird dafür plädiert, bei der Beschreibung dieser Vorgänge wissenschaftstheoretische Konzepte einzusetzen, die ursprünglich für die Behandlung empirisch-deskriptiver Theorien entwickelt worden sind. Das wird zunächst am Beispiel klassisch-utilitaristischer Theorien illustriert, für die Lakatos' einfache informelle Konzeption einen hinreichenden Beschreibungsrahmen abgibt. Anschließend wird skizziert, wie höher entwickelte metatheoretische Konzepte dazu eingesetzt werden können, um die Systematisierungsleistungen formal anspruchsvollerer ethischer Konzeptionen und ihrer empirischen Vortheorien zu erfassen.

4. Gründe für die Wahl klassisch-utilitaristischer Objekttheorien

Für die Wahl von Objekttheorien, die zum Paradigma des klassischen Utilitarismus gehören, sind drei Gründe maßgeblich. Der erste Grund ist philosophiehistorischer Natur: Der neuzeitliche Utilitarismus ist in einer Epoche — nämlich im ausgehenden 17. und 18. Jahrhundert — entstanden, in der die englische Moral-

[3] Dabei ist allerdings darauf zu achten, daß sich die Beobachtungsaussagen einerseits und die basalen Moralurteile andererseits auf einem vergleichbaren Allgemeinheitsniveau befinden: Häufig werden in unzulässiger Weise konkrete Beobachtungsaussagen mit wesentlich allgemeineren, vermeintlich basalen moralischen Urteilen verglichen. Für diesen Hinweis danke ich Herrn Dr. Hinsch. Zur Frage der Bedeutung der allgemeinen Akzeptierbarkeit derartiger Werturteile vgl. auch *Hackmann* (1976).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

philosophie unter dem Eindruck des überwältigenden Erfolges der Naturwissenschaften stand. Insbesondere Newtons Theorie galt nicht nur als Paradebeispiel einer erfolgreichen empirischen Theorie, sondern hatte auch Modellcharakter für den Entwurf ethischer Konzeptionen. Das kommt etwa in der folgenden Bemerkung Benthams zum Ausdruck:

> „Any . . . work of mine that has been or will be published on the subject of legislation or any other branch of moral science is an attempt to extend the experimental method of reasoning from the physical branch to the moral. What Bacon was to the physical world, Helvetius was to the moral. The moral world has therefore had its Bacon, but its Newton is yet to come." (*Bentham,* Economic Writings [1952-54, S. 100 f.])

Nahezu alle zum klassischen Utilitarismus gehörenden ethischen Konzeptionen wurden explizit nach dem Vorbild deskriptiver Theorien konstruiert. (Für Details vgl. *Gähde* [1992a, S. 65 ff.]) Es lag deswegen nahe, gerade bei diesen ethischen Theorien im Hinblick auf ihren Aufbau und ihre Funktionsweise eine besondere Nähe zu deskriptiv-empirischen Theorien zu vermuten.

Der zweite Grund besteht in einer charakteristischen Eigenschaft utilitaristischer Theorien, die sie grundlegend etwa von transzendentalphilosophischen Ansätzen in der Tradition Kants unterscheidet: Während jene Konzeptionen danach strebten, soweit wie möglich ohne empirische Annahmen auszukommen, die Ethik gerade *empiriefrei* zu halten, machte der klassische Utilitarismus bewußt und in besonderem Maße von empirischen Vortheorien Gebrauch. Die utilitaristische Ethik bietet damit die Möglichkeit, das Zusammenspiel deskriptiver und normativer Komponenten in ethischen Theorien im Detail zu studieren und seine Bedeutung für die Verteidigungsanstrengungen dieser Theorien zu analysieren.

Der dritte Grund schließlich ist der entscheidende: Die utilitaristische Ethik war von Anfang an weniger grundlagen- als vielmehr anwendungsorientiert. So standen etwa bei den Pionieren der utilitaristischen Bewegung, bei Bentham und Mill, rechtsformatorische, ökonomische und sozialpolitische Zielsetzungen im Vordergrund. Gerade im Bereich der angewandten Ethik liegen aber zugleich die deutlichsten moralischen Intuitionen vor, die mit konkreten normativen Implikationen des Utilitarismus entsprechend scharf kollidieren können. Gerade dort sind demnach auch die heftigsten Versuche zu erwarten, das Fundamentalprinzip der ethischen Theorie gegen den Vorwurf anti-intuitiver Implikationen zu verteidigen.

Tatsächlich sah sich die utilitaristische Ethik seit ihren ersten Anfängen stets heftigen Angriffen ausgesetzt. Insbesondere wurde ihr das Fehlen eines eigenen normativen Konzepts vorgeworfen, das die Verteilung des Gesamtnutzens auf die betroffenen Individuen steuern könnte. Dieser Mangel der utilitaristischen Theorie erzeugt eine Vielzahl von Kollisionen mit basalen moralischen Intuitionen. Im ökonomischen Bereich führt das Fehlen eines eigenen Verteilungsprinzips dazu, daß, zumindest im Prinzip, der Zustand des kollektiven Nutzenmaximums

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ein Zustand extremer sozialer Ungleichheit sein kann. Weiterhin führt es zu einer nach herkömmlichem Rechtsempfinden inakzeptablen juristischen Konsequenz: Die Maximierung des Gesamtnutzens kann — unter bestimmten Bedingungen — die Bestrafung Unschuldiger erfordern. Diese Konsequenz wurde bereits 1758 von C. A. Helvétius in aller Schärfe gesehen (und im übrigen auch befürwortet!) (*Helvétius* Œuvres Complètes [1795, 1967-69, Bd. I/II, S. 80 ff.]). Dadurch wurde für den Utilitarismus die Sicherung von Grundrechten zu einem notorischen und kaum lösbaren Problem (vgl. *Dworkin* [1977], *Frey* [1985]).

Charakteristisch für diese Einwände ist, daß sie, einmal erhoben, nie mehr aus der Diskussion verschwunden sind. Anti-intuitive Implikationen wirken analog wie *Anomalien* im Bereich empirisch-deskriptiver Theorien: Sie stellen eine ständige, andauernde Herausforderung der Theorie dar und senken deren Akzeptanz. Von seiten der Gegner führen sie zu hartnäckig vorgebrachten Angriffen, von seiten der Verfechter zu ebenso hartnäckigen Verteidigungsanstrengungen in Form von Modifikationen der Theorie.

Die utilitaristische Ethik hat diese Angriffe nicht nur bis heute überlebt, sondern stellt insbesondere im angelsächsischen Sprachraum nach wie vor eine der prominentesten moralphilosophischen Positionen dar. Es liegt deswegen nahe, gerade bei diesen ethischen Objekttheorien danach zu fragen, worauf ihre außerordentliche Widerstandskraft gegen Konflikte mit grundlegenden moralischen Intuitionen beruht. Dieser Frage wenden wir uns nun zu. Die Grundidee der folgenden Überlegungen besteht darin, daß utilitaristische Theorien Mechanismen zu ihrer Verteidigung einsetzen, die weitgehend den zuvor geschilderten Immunisierungsmechanismen deskriptiver Theorien ähneln.

5. Der harte Kern klassisch-utilitaristischer Theorien

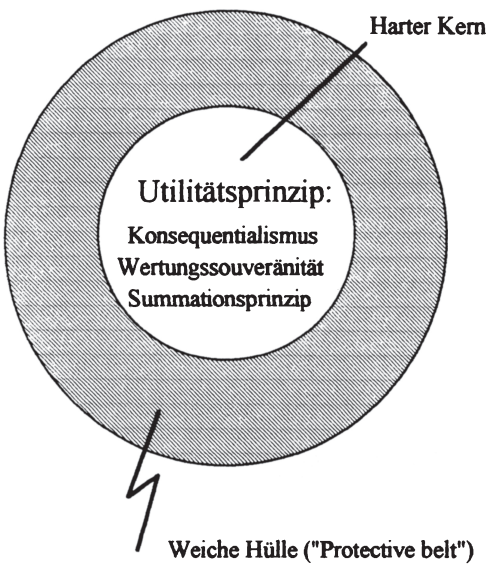
Um Lakatos' *Methodologie der Forschungsprogramme* auf utilitaristische Theorien anwenden zu können, muß zunächst der *harte Kern* dieser Theorienfolge identifiziert werden. Im Fall des klassischen Utilitarismus besteht dieser harte Kern im wesentlichen aus dem Utilitätsprinzip: einem Fundamentalprinzip, das von Benthams Pionierwerk *An Introduction to the Principles of Morals and Legislation* (1789) über Mills Essay *Utilitarianism* (1861) bis hin zu Sidgwicks *Methods of Ethics* (1874) nahezu unverändert vertreten worden ist. Dieses Prinzip kann — einem Vorschlag von *Sen* (1979) folgend — in drei informelle Kernthesen aufgespalten werden:

1. Handlungen werden ausschließlich nach ihren Konsequenzen beurteilt *(Konsequentialismus)*.

2. Der Nutzen von Handlungskonsequenzen wird *ausschließlich* nach den Wertmaßstäben der von diesen Konsequenzen betroffenen Individuen beurteilt *(Wertungssouveränität)*.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

3. Der individuelle Gesamtnutzen einer Handlung wird berechnet als Summe der von dem betreffenden Individuum ermittelten Nutzenwerte aller Handlungskonsequenzen. Der kollektive Gesamtnutzen wird erhalten, indem alle individuellen Nutzenwerte interpersonell aufaddiert werden. Moralisch geboten ist eine Handlung genau dann, wenn sie den kollektiven Gesamtnutzen stärker erhöht als jede andere mögliche Handlung *(Summations-Forderung).*



Enthält Theorieteile zur Interpretationen des Nutzen- und Individuenbegriffs

Wir werden im folgenden zunächst auf diese *informelle* Charakterisierung der Kernthesen des klassischen Utilitarismus Bezug nehmen. In Abschnitt 8 wird ein wissenschaftstheoretischer Rahmen skizziert, in dem auch *formale Präzisierungen* dieser Kernthesen dargestellt und wissenschaftstheoretisch analysiert werden können.

6. Verteidungsmechanismen und Begriffsentwicklung

In der Wissenschaftstheorie herrscht weitgehend Einigkeit darüber, daß die Fruchtbarkeit zur Systematisierung einer Datenbasis *das* entscheidende Kriterium darstellt, das die Begriffsentwicklung einer empirischen Theorie steuert: Die Rolle, die ein Begriff in einer deskriptiven Theorie spielt, hängt wesentlich von

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

der Gültigkeit und dem Gewicht empirischer Hyothesen ab, die sich mit seiner Hilfe formulieren lassen.[4]

Bei normativ-ethischen Theorien scheint die Situation auf den ersten Blick grundlegend anders zu sein. Eine Beobachtungsbasis zur Beurteilung der Gültigkeit normativer Prinzipien gibt es nicht. Damit fällt nach Auffassung der meisten Philosophen auch die Fruchtbarkeit zur empirischen Hypothesenbildung als zentrales Steuerungsorgan für die Begriffsentwicklung fort: Ethische Theorien müssen nach dieser Auffassung die Bereitstellung und Entwicklung ihres begrifflichen Apparats aus anderen Quellen speisen als deskriptive Theorien.[5]

In den folgenden Abschnitten wird für eine Sichtweise argumentiert, die in scharfem Gegensatz zu dieser Standardauffassung steht. Danach stellt auch bei ethischen Theorien eine Systematisierungsleistung das entscheidende Steuerungsorgan für die Begriffsentwicklung dar — in weitgehender Analogie zur Begriffsentwicklung bei deskriptiven Theorien. Dabei handelt es sich um die Systematisierungsleistung, die diese Theorien gegenüber basalen oder für basal gehaltenen moralischen Intuitionen erbringen. Diese These wird im folgenden am Beispiel zweier Grundbegriffe der klassischen utilitaristischen Ethik erläutert: der Begriffe *betroffene Individuen* bzw. *individueller Nutzen* (Abschnitte 6.1 bzw. 6.2).

Beide Grundterme treten in den zuvor genannten Kernthesen wesentlich auf. Ihre Bedeutung bleibt aber in diesen Kernthesen selbst unbestimmt. Bei allen Autoren des klassischen Utilitarismus wird die Bedeutung dieser Begriffe jedoch durch ergänzende, von Autor zu Autor unterschiedliche Festlegungen näher eingegrenzt. Im Fall des Individuenbegriffs handelt es sich dabei im allgemeinen um einfache Bedeutungszuweisungen. Dagegen wird die Bedeutung des Nutzenbegriffs durch eigene *Subtheorien* näher bestimmt, in denen zum Teil in erheblichem Maß auf empirische (insbesondere ökonomische und motivationspsychologische) Vortheorien Bezug genommen wird.

Da bereits in der klassischen Entwicklungsphase sowohl die Interpretationen des Individuen- als auch die des Nutzenbegriffs variieren, sind sie — in Lakatos' Bild — nicht mehr dem *harten Kern* des Utilitarismus zuzurechnen. Sie bilden vielmehr Bestandteile seiner *weichen Hülle*. In den folgenden beiden Unterabschnitten wird gezeigt, daß und in welcher Weise sie zudem die Funktion eines *protective belt* für das klassisch-utilitaristische Programm übernehmen.

Im folgenden Unterabschnitt wird zunächst auf die wechselnden Bedeutungszuweisungen zum Begriff der *betroffenen Individuen* eingegangen. Gerade ein

---

[4] Ein Beispiel, das diesen Sachverhalt illustriert, ist die herausragende Rolle von Erhaltungsgrößen in naturwissenschaftlichen (insbesondere physikalischen) Theorien.

[5] So vertritt etwa Richard Hare die Auffassung, daß die sprachphilosophische Analyse linguistischer Intuitionen das begriffliche Baumaterial zur Verfügung stellen soll, aus dem ethische Theorien zu konstruieren sind; vgl. *Hare* (1988).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Blick in die — für Ökonomen gewiß exotisch anmutende — „Frühgeschichte" des Utilitarismus macht deutlich, in welchem Maße Bedeutungsverschiebungen dieses Grundterms für den Zweck einer kohärenten Systematisierung vorgegebener (und wechselnder) moralischer Intuitionen eingesetzt wurden.

6.1 Zur Entwicklung des Individuenbegriffs im klassischen Utilitarismus

Wie problematisch eine Interpretation des Individuenbegriffs für die utilitaristische Ethik sein kann, zeigt sich bereits bei einem der frühesten Pioniere dieser Position, bei Richard Cumberland. In *De Legibus Naturae* (in der er die Naturrechtslehre des Hugo Grotius gegen Hobbes verteidigt) spricht er 1672 in auffallender Weise nicht von Menschen, deren Nutzen zu maximieren sei, sondern von *rational beings.* Cumberlands weitere Ausführungen verraten, daß er unter diesen Begriff menschliche Individuen *und Gott* subsumiert (vgl. dazu die Einleitung zu *Cumberland* [1672] in der engl. Übersetzung von 1727). Er entwickelt damit eine Variante des sog. *theologischen Utilitarismus,* wie er wenig später auch in *Paley* (1785, dt. 1986) vertreten worden ist. Ziel dieser Interpretation des Individuenbegriffs ist es, die normativen Implikationen des Utilitarismus mit vorgegebenen religiösen Intuitionen zu versöhnen. Dabei sollte angefügt werden, daß Cumberland, wie viele andere Hauptvertreter der englischen Ethik, Theologe war — nämlich anglikanischer Bischof von Peterborough / Grafschaft Kent.

In der ersten systematischen Darstellung der utilitaristischen Ethik, in Benthams *Introduction to the Principles of Morals and Legislation* (1789), wird dagegen bereits eine säkulare Interpretation des Individuenbegriffs vorgenommen. Bentham spricht von *sentient beings,* von *empfindungsfähigen Wesen.* Dadurch eröffnet sich ihm die Möglichkeit, Tiere im Nutzenkalkül zu berücksichtigen. Ausgehend von der zuvor dargestellten hedonistischen Interpretation des Nutzenbegriffs fragt Bentham in einer Fußnote im 17. Kapitel seiner *Principles:*

> „What else is it that should trace the insuperable line? Is it the faculty of reason, or, perhaps, the faculty of discourse? But a full-grown horse or dog, is beyond comparison a more rational, as well as a more conversible animal, than an infant of a day, or a week, or even a month, old. But suppose the case were otherwise, what would it avail? the question is not, Can they *reason?* nor, Can they *talk?* but, Can they *suffer?*" (*Bentham* [1789, S. 283n]).

Bentham beschränkt sich nicht auf die Forderung, Tiere zum Objekt der moralischen Fürsorge zu machen. Er fordert darüber hinaus die strafrechtliche Verankerung eines Verbotes von Grausamkeit gegenüber Tieren. Dabei zieht er eine Parallele zur Frage der Sklaverei:

> „The day has been, I grieve to say in many places it is not yet past, in which the greater part of the species, under the denomination of slaves, have been treated by the law exactly upon the same footing, as, in England for example, the inferior races of animals are still. The day may come, when the rest of the animal creation

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

> may acquire those rights which never could have been withholden from them but by the hand of tyranny." (a. a. O.).

Es ist hier nicht der Ort, um auf Details und Schwierigkeiten der Benthamschen Position einzugehen. Vielmehr kommt es im Rahmen der vorliegenden Argumentation nur darauf an, den folgenden Sachverhalt zu illustrieren: Bentham nimmt gegenüber Cumberland eine grundlegende Uminterpretation des Individuenbegriffs vor. Indem er nicht mehr von *rational beings,* sondern statt dessen von *sentient beings* spricht, gelingt es ihm, wesentlich andersartige moralische Intuitionen — hier zu moralischen Fragen der Behandlung von Tieren — in seine utilitaristische Position zu integrieren. Das Ziel der Systematisierung vorgegebener moralischer Intuitionen steuert hier offensichtlich die Entwicklung des begrifflichen Apparates seiner utilitaristischen Position.

Die Auflistung von Uminterpretationen des Individuenbegriffs im Rahmen utilitaristischer Theorien läßt sich nahezu beliebig fortsetzen. So spielte etwa im viktorianischen Zeitalter die Frage der Berücksichtigungfähigkeit und -würdigkeit der Ureinwohner der englischen Kolonien im utilitaristischen Kalkül eine zentrale Rolle (vgl. *Pollard* [1992]).

Eine weitere einschneidende Veränderung des Individuenbegriffs, die insbesondere für die Diskussion in der angewandten Ethik von grundlegender Bedeutung war, vollzog sich in der zweiten Hälfte des neunzehnten Jahrhunderts: Der Individuenbegriff wurde so erweitert, daß auch *zukünftige* menschliche Individuen einbezogen werden konnten. Zugleich wurden die Konsequenzen dieser Erweiterung für den Nutzenbegriff diskutiert. Damit begann die *future-generations*-Debatte im Rahmen der utilitaristischen Ethik (*Sikora / Barry* [1978]). Eine der frühesten expliziten Bemerkungen zu diesem Thema findet sich in *Sidgwick* (1874, [7]1981, S. 414):

> „In the first place, it may be asked, How far we are to consider the interests of posteriority when they seem to conflict with those of existing human beings? It seems, however, clear that the time at which a man exists cannot affect the value of his happiness from a universal point of view; and that the interest of posterity must concern a Utilitarian as much as those of his contemporaries, except in so far as the effect of his actions on posterity — and even the existence of human beings to be affected — must necessarily be more uncertain."

Sidgwick ist sich auch bereits der Schwierigkeiten bewußt, die daraus resultieren, daß die Anzahl der im hedonistischen Kalkül zu berücksichtigenden Individuen zu einem steuerbaren und damit selbst nach dem Utilitätsprinzip zu bewertenden Parameter wird. Dabei wird deutlich, daß das Utilitätsprinzip zu außerordentlich problematischen Konsequenzen in bezug auf die Begrenzung des Bevölkerungswachstums führt:

> „For if we take Utilitarianism to prescribe, as the ultimate end of action, happiness on the whole, and not any individual's happiness, unless considered as an element of the whole, it would follow that, if the additional population enjoy on the whole

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

positive happiness, we ought to weigh the amount of happiness gained by the extra number against the amount lost by the remainder. So that, strictly conceived, the point up to which, on Utilitarian principles, population ought to be encouraged to increase, is not that at which average happiness is the greatest possible — as appears to be often assumed by political economists of the school of Malthus — but that at which the product formed by multiplying the number of persons living into the amount of average happiness reaches its maximum." (*Sigdwick* [1874, [7]1981, S. 415 f.])

Sidgwick diskutiert nicht nur die Alternative *Durchnittsnutzen-Utilitarismus* versus *Nutzensummen-Utilitarismus,* sondern behandelt auch die im klassischen Utilitarismus des neunzehnten Jahrhunderts teilweise vernachlässigte Problematik alternativer Verteilungsprinzipien für eine vorgegebene Gütermenge. Auf Details dieser Analyse kann an dieser Stelle nicht näher eingegangen werden. Entscheidend für den hier behandelten Problemkontext ist allein die folgende Tatsache: Sidgwick versucht den Individuenbegriff so zu erweitern, daß dieser auch moralischen Intuitionen in bezug auf die Behandlung zukünftiger Generationen gerecht werden kann. Diese Erweiterung hat offensichtlich zugleich Konsequenzen in bezug auf mögliche Interpretationen des Nutzenbegriffs (vgl. Abschnitt 8).

Diese Beispiele bestätigen zunächst die zentrale These von Abschnitt 3: Die klassisch-utilitaristischen Theorien sprechen basalen moralischen Urteilen im allgemeinen explizit eine gewisse Korrektivkraft für die ethische Theorienbildung zu. Konflikte zwischen normativen Implikationen des Utilitarismus und vorgegebenen moralischen Intuitionen werden in aller Regel nicht einfach hingenommen. Statt dessen führen sie dazu, daß die utilitaristische Theorie *und* die konfligierenden moralischen Intuitionen überprüft werden. Wird dieser Konflikt zu Lasten der Theorie entschieden, so wird die erforderliche Korrektur in einer Weise vorgenommen, bei der am Utilitätsprinzip selbst unverändert festgehalten werden kann. In den zuvor besprochenen Fällen wurde das durch mehrfache Uminterpretationen des Begriffs der *betroffenen Individuen* ermöglicht. Die Fruchtbarkeit zur Systematisierung vorgegebener moralischer Intuitionen stellt, wie die Beispiele illustrieren, zumindest *eines* der entscheidenden Kriterien dar, die die Begriffsentwicklung ethischer Termini im Utilitarismus gesteuert haben.

6.2 Bemerkungen zur Entwicklung des Nutzenbegriffs im klassischen Utilitarismus

Ein zweites Kriterium wird deutlich, wenn man die mehrfachen Bedeutungsverschiebungen betrachtet, die der Nutzenbegriff im klassischen Utilitarismus erfahren hat: Operationalisierungs- und Meßprobleme stellen bereits in der klassischen Entwicklungsphase der utilitaristischen Ethik ein zweites, wichtiges Steuerungskriterium für die Begriffsentwicklung dieser Theorien dar — in weitgehender

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Analogie zu der Bedeutung, die dieses Kriterium für die konzeptuelle Entwicklung empirisch-deskriptiver Theorien besitzt. Die Interpretationen des Nutzenbegriffs unterscheiden sich dabei von den bloßen Bedeutungszuweisungen zum Individuenbegriff daduch, daß sie im allgemeinen im Rahmen eigener, teilweise recht ausgedehnter peripherer Theorieteile vorgenommen werden. Ich habe den Wandel des Nutzenbegriffs an anderer Stelle ausführlich dargestellt und werde mich in den folgenden Darstellungen auf die Rolle konzentrieren, die empirische Hypothesen im Rahmen dieses Prozesses gespielt haben.[6]

In der von Bentham vorgelegten Fassung des Utilitarismus wurde der Nutzenbegriff im Rahmen einer hedonistischen Position interpretiert. Diese Position schloß zwei grundlegend verschiedene Aspekte des Hedonismus ein. Der erste Aspekt bestand in einer deskriptiven psychologischen Theorie, nach der menschliches Verhalten vollständig durch ein einziges deterministisches Gesetz gesteuert wird: durch die Suche nach Lust sowie den Versuch der Vermeidung von Schmerzen. Dieser psychologisch-deskriptive Aspekt des Hedonismus wurde durch einen normativ-ethischen Aspekt ergänzt: Der Nutzen, den ein Sachverhalt (etwa eine bestimmte Handlungskonsequenz) für ein konkretes Individuum besitzt, besteht danach in einer Gefühlsbilanz, nämlich der Menge der von diesem Sachverhalt hervorgerufenen Lust abzüglich der entsprechenden Menge an Schmerz (vgl. *Bentham* [1789, 1970], insbesondere Kap. I, IV). Das Utilitätsprinzip fordert, diese Gefühlsbilanz interpersonell zu optimieren.

Offensichtlich ist die Kombination von psychologischem Hedonismus einerseits und ethischem Hedonismus andererseits mit erheblichen Schwierigkeiten belastet. Wenn Individuen ohnehin die Maximierung ihrer Lustbilanz anstreben, wieso muß dies dann noch im Rahmen einer normativ-ethischen Theorie ausdrücklich gefordert werden? Benthams Antwort: Individuen streben zwar nach der Maximierung ihrer *individuellen* Empfindungsbilanz; das Utilitätsprinzip fordert jedoch etwas grundlegend anderes, nämlich die Maximierung des *kollektiven* Gesamtnutzens. Damit drängt sich aber unmittelbar die nächste Frage auf: Wenn menschliches Handeln durch ein deskriptives deterministisches Gesetz (das das Streben nach Maximierung des individuellen Nutzens zur einzigen Triebfeder menschlichen Handelns erklärt) bereits vollständig festgelegt ist, wie kann dann eine normativ-ethische Theorie etwas anderes (nämlich die Maximierung des kollektiven Gesamtnutzens) fordern? Benthams Antwort: Diese Forderung richtet sich nicht an einzelne Individuen. Vielmehr ist sie als eine Bedingung zu interpretieren, die jede sinnvolle Gesetzgebung zu erfüllen hat. Eine prästabilierte Harmonie zwischen individueller und kollektiver Nutzenmaximierung existiert nicht. Die Aufgabe der Gesetzgebung besteht darin, diese Harmonie mit

---

[6] Eine ausführlichere Darstellung der Entwicklung des Nutzenbegriffs von Benthams *Introduction to the Principles of Morals and Legislation* bis Sidgwicks *Methods of Ethics* findet sich in *Gähde* (1992 a). Der vorliegende Unterabschnitt enthält eine Zusammenfassung der dortigen Ausführungen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Hilfe staatlicher Sanktionen künstlich zu erzwingen. Die Gesetzgebung (und insbesondere die Strafgesetzgebung) ist so vorzunehmen, daß individuelle und kollektive Nutzenmaximierung koinzidieren.

Um diese Aufgabe erfolgreich bewältigen zu können, müssen im hedonistischen Utilitarismus die Empfindungsbilanzen bestimmt werden, die Individuen als Ergebnis der Ausführung ihrer Handlungen erwarten. Diese (antizipierten) Empfindungsbilanzen beruhen im wesentlichen auf Erfahrungen, die die Individuen in der Vergangenheit bei der Ausführung ähnlicher Handlungen bereits gemacht haben. Der Versuch einer praktischen Anwendung des hedonistischen Utilitarismus führt damit zu einer anspruchsvollen empirischen Aufgabenstellung: Es müssen Verfahren entwickelt werden, mit denen das Maß von Lust- oder Schmerzempfindungen (in Benthams Terminologie: „measure of sentiment") bestimmt werden kann, die sich bei einem Individuum als Folge der Ausführung bestimmter Handlungen einstellen, bzw. die es als Ergebnis der Ausführung zukünftiger Handlungen erwartet.

Wissenschaftshistorische Untersuchungen haben in den vergangenen Jahrzehnten zahlreiche neue Quellen zum Utilitarismus erschlossen. Dadurch ist insbesondere deutlich geworden, in welchem Ausmaß bereits die klassischen Utilitaristen diese erfahrungswissenschaftliche Aufgabe ernst genommen haben. Vor allem Bentham hat sich dabei, wie *McReynolds* (1968) nachgewiesen hat, weit auf das Gebiet der experimentellen Psychologie (insbesondere der Motivationspsychologie) vorgewagt. Im Anschluß an Arbeiten des französischen Physikers und Mathematikers Pierre Maupertuis[7] hat er zwei verschiedene Verfahren zur Messung von Empfindungsstärken *(measure of sentiment)* vorgeschlagen. Beim ersten, *introspektiven* Verfahren werden zunächst Empfindungsintensitäten als Vielfache einer gerade noch wahrnehmbaren Grenzintensität abgeschätzt. Daraus wird anschließend das Empfindungsmaß, in das neben der Intensität auch die Dauer der Empfindung sowie einige weitere Parameter eingehen, ermittelt (*Gähde* [1992a, S. 94 ff.]). Diese Vorgehensweise kann als ein Vorläufer von Schwellenwertverfahren angesehen werden, die später in der experimentellen Psychophysik und im Rahmen der psychologischen Skalierungsverfahren eine erhebliche Bedeutung erlangt haben (*Fechner* [1860, Kap. X und XI]). Bei dem zweiten, *behavioristischen* Verfahren wird dagegen die von einem Individuum erwartete Empfindungsbilanz direkt aus seinem Entscheidungsverhalten bestimmt (vgl. *Gähde* [1992a, S. 96]).

Wichtig ist, daß in beiden von Bentham entworfenenen Verfahren zur Messung hedonistisch interpretierter Nutzenwerte ausschließlich auf quantitativ faßbare Aspekte von Empfindungen Bezug genommen wurde. Bentham war sich der qualitativen Unterschiede zwischen verschiedenen Empfindungsformen durchaus

---

[7] Vgl. dazu Maupertuis „Essai de philosophie morale" (1749) in: *Maupertuis*, Œuvres (1768; 1758, Bd. 1)

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

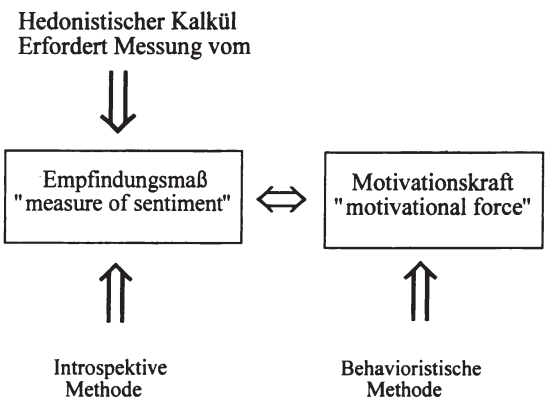
bewußt (vgl. etwa *Bentham* [1789, Kap. V]); er bestritt jedoch deren moralische Relevanz. Das hatte eine stark anti-intuitive Konsequenz seines hedonistischen Kalküls zur Folge: Bentham war gezwungen, beliebige Lustempfindungen als positive Beiträge zur Nutzenbilanz einer Handlung zu berücksichtigen — darunter auch solche Empfindungen, die nach herkömmlicher moralischer Intuition eindeutig abzulehnen waren. Daß etwa ausgerechnet das Vergnügen, das eine Person aus der Schädigung anderer Personen bezieht, den moralischen Wert einer Handlung erhöhen sollte, war nicht nur für Benthams Zeitgenossen intuitiv kaum nachzuvollziehen. Diese Konsequenz wurde vielmehr als entscheidendes Defizit — als *Anomalie* — seiner utilitaristischen Theorie betrachtet.

Bereits der zweite Hauptvertreter des Utilitarismus im neunzehnten Jahrhundert, J. St. Mill, versuchte deswegen die utilitaristische Theorie so umzubauen, daß gleichzeitig zwei Ziele erreicht wurden: Auf der einen Seite wollte er die anti-intuitive Konsequenz der Benthamschen Konzeption vermeiden. Auf der anderen Seite wollte er am Utilitätsprinzip — als dem Fundamentalprinzip der Theorie — unbedingt festhalten. Mill gab dazu Benthams *quantitativ-hedonistische* Subtheorie zur Interpretation des Nutzenbegriffs auf und ersetzte sie durch eine *qualitativ-hedonistische* Subtheorie. Der entscheidende Unterschied bestand darin, daß nun verschiedenen Empfindungsformen unterschiedliche moralische Wertigkeiten zugeordnet wurden. Als Grundlage für die Abwägung von Qualität und Quantität verschiedener Empfindungsarten sollte das beobachtete Wahlverhalten *kompetenter Personen* in geeignet modellierten Entscheidungssituationen gewählt werden. Als *kompetent* betrachtete Mill dabei Individuen, die auf Grund ihrer Geburt oder sozialen Stellung in der Lage gewesen waren, ihre intellektuellen und emotionalen Fähigkeiten auszubilden und die ganze Breite möglicher Empfindungsformen kennenzulernen. Von diesen Individuen erwartete Mill, daß sie nicht nur zu einer Abwägung zwischen verschiedenen Empfindungsstärken, sondern auch zwischen verschiedenen Empfindungsformen befähigt wären. Hinter diesem — keineswegs völlig überzeugenden — Vorschlag verbirgt sich ein recht modern anmutendes Dilemma: Mill wollte das utilitaristische Grundprinzip der Wertungssouveränität aufrechterhalten und damit jede direkte paternalistische Bevormundung von Individuen vermeiden, ohne deswegen zugleich beliebige Bedürfnisse und Empfindungen ohne jede weitere moralische Wertung im utilitaristischen Kalkül akzeptieren zu müssen.

Durch die Ersetzung der quantitativ-hedonistischen durch eine qualitativ-hedonistische Interpretation des Nutzenbegriffs erfuhr dieser einen grundlegenden Bedeutungswandel. Dadurch wurde zugleich eine Bedeutungsverschiebung des Utilitätsprinzips selbst induziert: Die Klassen der mit diesem Prinzip konformen bzw. konfligierenden Handlungen änderten sich. Dabei konnte das Utilitätsprinzip seine äußere Gestalt jedoch unverändert beibehalten: Es wurde auf Kosten peripherer Theorieteile gegen den Einwand anti-intuitiver Implikationen verteidigt.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Im hedonistischen Utilitarismus war zunächst angenommen worden, daß das *Maß* einer Empfindung (im Sinne der Benthamschen Definition) und ihre handlungsauslösende *(motivationale)* Kraft im allgemeinen zusammenfallen. Nur unter dieser Annahme konnte davon ausgegangen werden, daß die beiden von Bentham vorgeschlagenen Verfahren zur Nutzenmessung — das *introspektive* und das *behavioristische* Verfahren — zum selben Ergebnis führen. Mit fortschreitender Entwicklung der Psychologie im 19. Jahrhunderts wurde zunehmend deutlich, daß diese Voraussetzung empirisch unhaltbar ist. Die Entkoppelung der beiden Konzepte hatte eine fatale Konsequenz für die hedonistische Variante des Utilitarismus: Sie schnitt den behavioristischen Zugang zur Nutzenmessung ab und lieferte die Utilitaristen so ausschließlich dem introspektiven Verfahren aus. Der Versuch der Entwicklung leistungsfähiger Meßverfahren zur *introspektiven* quantitativen Erfassung von Empfindungen konfrontierte die Utilitaristen mit erfahrungswissenschaftlichen Problemen, die sie nicht bewältigen konnten und die schließlich erheblich zum Zusammenbruch der hedonistischen Spielart des Utilitarismus beitrugen (*Gähde* [1992a, S. 104-108]).



In dem Maße, in dem die Undurchführbarkeit eines hinreichend präzisen introspektiven Verfahrens zur Nutzenmessung deutlich wurde, sank die Attraktivität der hedonistischen Varianten des Utilitarismus. Zugleich wurde durch diese Entwicklung die Entstehung eines neuartigen Typs utilitaristischer Theorien gefördert. Bei diesen Theorien wird der Nutzenbegriff nicht mehr über *Empfindungsstärken,* sondern über die *motivationale Kraft* antizipierter Empfindungen interpretiert (die wiederum aus dem Entscheidungsverhalten von Individuen erschlossen werden kann). Der Übergang vom hedonistischen Utilitarismus zum Präferenzutilitarismus wurde so weniger durch genuin philosophische Überlegungen — etwa zur Gestalt des *summum bonum* —, als vielmehr durch erfahrungswissenschaftliche Probleme bei der Nutzenmessung vorbereitet. Dieser Übergang

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

erwies sich als von grundlegender Bedeutung nicht nur für die weitere Entwicklung des utilitaristischen Programms selbst, sondern auch für die Entstehung der modernen wohlfahrtsökonomischen Theorien.

7. Verteidungsmechanismen und empirische Vortheorien

Wie bereits zuvor erwähnt, bieten sich klassisch-utilitaristische Konzeptionen auch deswegen für einen Vergleich mit deskriptiven Theorien besonders an, weil sie bewußt und in erheblichem Maß von erfahrungswissenschaftlichen Hypothesen Gebrauch machten. Welche Rolle derartige empirische Behauptungen im Zusammenhang mit Operationalisierungs- und Meßproblemen spielten, dürften bereits die Ausführungen zur hedonistischen Nutzenmessung im letzten Abschnitt gezeigt haben. Daneben wurden erfahrungswissenschaftliche Hypothesen aber auch gezielt eingesetzt bei dem Versuch, das Utilitätsprinzip zu begründen bzw. gegen Konflikte mit vorgegebenen normativen Zielvorstellungen zu verteidigen (vgl. *Gähde* [1993b]). Das läßt sich gerade am Beispiel der utilitaristischen Einstellung zu Fragen der distributiven Gerechtigkeit verdeutlichen.[8]

Eine besondere Bedeutung kam dabei dem *Gesetz des abnehmenden Grenznutzens* zu (vgl. *Gossen* [1854, ³1927]). Vorläufer dieses Gesetzes wurden bereits zu Benthams Zeiten diskutiert und waren den Utilitaristen auch bekannt. Bentham selbst betont den abnehmenden Grenznutzen des Geldes ausdrücklich. In einem kleineren, von seinem Genfer Freund Etienne Dumont zunächst in französischer Sprache herausgegebenen Werk mit dem Titel *Traités de législation civile et pénale* (1802) schreibt er (ich zitiere die englische Übersetzung): „the excess in happiness of the richer will not be so great as the excess of his wealth".[9] Darüber hinaus zieht er die normative Konsequenz aus dem Gesetz des abnehmenden Grenznutzens und dem Utilitätsprinzip, indem er Gleichheit als einen wesentlichen Faktor menschlichen Glücks bezeichnet. Dennoch lehnt er — ebenso wie später etwa James Mill (*Bain* [1882], S. 364-5) — alle Forderungen nach einer grundlegenden Umverteilung gesellschaftlicher Güter und Chancen kategorisch ab. Seine Argumentation kann wie folgt zusammengefaßt werden: Die utilitaristisch begründete Forderung nach Gleichheit kollidiert mit der ebenfalls utilitaristisch begründeten Forderung nach Sicherheit des Eigentums. Dabei wird, so Bentham, der durch die Zunahme der sozialen Gleichheit erzeugte Nutzengewinn überkompensiert durch die Nutzeneinbuße in Form einer Abnahme an Eigentumssicherheit — und, damit verbunden, der Abnahme der Bereitschaft, überhaupt Eigentum zu schaffen und so die wirtschaftliche Tätigkeit anzuregen.

---

[8] Für eine ausführliche Darstellung dieser Problematik aus wirtschafts- bzw. sozialhistorischer Sicht vgl. *Pollard* (1993, S. 20 ff.).

[9] Eine englische Übersetzung mit dem Titel „Principles of the Civil Code" ist im zweiten Band der von John Bowring herausgegebenen elfbändigen Ausgabe der Werke Benthams enthalten; vgl. *Bentham* Works (1838-43).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Diese Behauptung stellt im Rahmen der von Bentham vertretenen hedonistischen Interpretation des Nutzenbegriffs eine *erfahrungswissenschaftliche Hypothese* dar. Die Gültigkeit dieser Hypothese mußte gerade in Anbetracht der krassen Einkommens- und Besitzunterschiede im England des 18. und 19. Jahrhunderts als zweifelhaft erscheinen. Empirische Untersuchungen, die diese Behauptung hätten stützen können und die Benthams eigene Ansprüche an die Präzision der hedonistischen Nutzenmessung auch nur entfernt widerspiegeln würden, wurden von ihm nicht angestrengt. Der Verdacht liegt nahe, daß hier spekulative, im Graubereich der erfahrungswissenschaftlichen Erkenntnis liegende empirische Behauptungen eingesetzt wurden, um als Puffer zwischen dem Utilitätsprinzip und vorgegebenen normativen Zielsetzungen zu dienen.

8. Ausblick: Theorien-Netze und die utilitaristischen / wohlfahrtsökonomischen Theorien des zwanzigsten Jahrhunderts

Für eine wissenschaftstheoretische Analyse der einfachen, im allgemeinen umgangssprachlich formulierten utilitaristischen Theorien des neunzehnten Jahrhunderts reicht Lakatos' informelle wissenschaftstheoretische Konzeption aus. Dagegen bietet es sich an, für die Analyse der Nachfolgetheorien des Utilitarismus, wie sie im zwanzigsten Jahrhundert etwa im Rahmen der modernen Wohlfahrtsökonomie entstanden sind, anspruchsvollere wissenschaftstheoretische Konzepte einzusetzen. Das gilt insbesondere dann, wenn diese Objekttheorien in mathematisierter Gestalt vorliegen.

Im folgenden wird auf das Konzept der sog. *Theorien-Netze* eingegangen. Dieses Konzept wurde innerhalb einer bereits zuvor erwähnten speziellen Variante der semantischen Theorienauffassung, dem *strukturalistischen Ansatz,* entwikkelt. Dieser Ansatz ist, wie die anderen semantischen Metatheorien, durch den Einsatz modelltheoretischer Hilfsmittel gekennzeichnet. (Für Einzelheiten vgl. *Balzer / Moulines / Sneed* [1987]). Der Einfachheit halber werden das Konzept und seine Vorzüge im folgenden an Hand der bereits zuvor dargestellten Objekttheorien erläutert.

Ein *Theorien-Netz* besteht aus einem sog. *Basiselement* und einer größeren Anzahl *spezialisierter Theorie-Elemente.* Das Basiselement repräsentiert — anschaulich gesprochen — den größten gemeinsamen Nenner aller im Netz vertretenen speziellen Positionen. Es enthält die Grundprinzipien der Theorie. Dabei wird zwischen sog. *Charakterisierungen* und den *eigentlichen Axiomen* unterschieden. *Charakterisierungen* erfassen die formalen (mathematischen) Eigenschaften der Terme, die in der betreffenden Theorie auftreten. Wie wichtig diese Informationen für ethische Theorien werden können, zeigt gerade das Beispiel utilitaristischer Konzeptionen. So kann im Rahmen von Charakterisierungen etwa festgelegt werden, auf welchem Skalenniveau die Nutzenfunktion zu messen ist

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

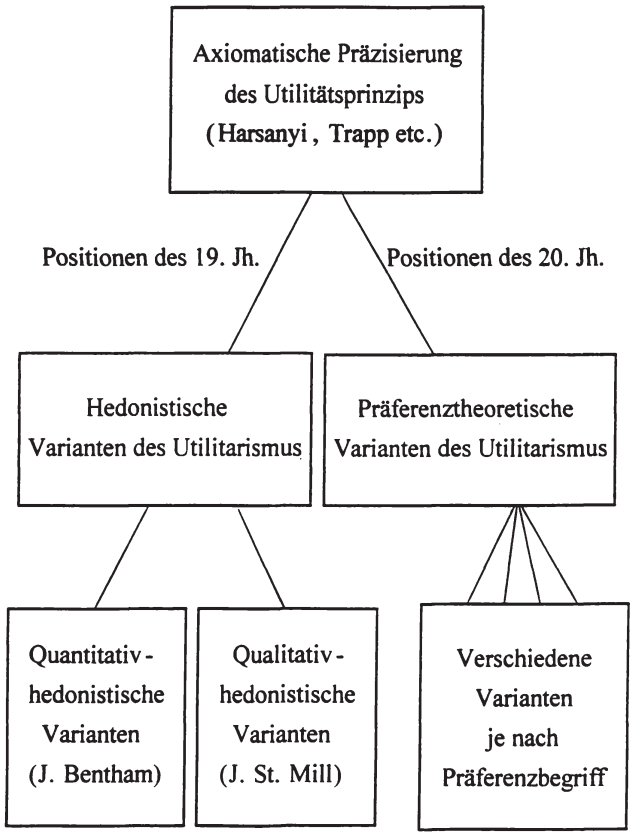
— eine Präzisierung, durch die, wie die Geschichte des Utilitarismus zeigt, zahllose Verwirrungen und Fehldeutungen hätten vermieden werden können. Ergänzend dazu erfassen die *eigentlichen Axiome* die Grundprinzipien der Theorie, im Fall der utilitaristischen Ethik also geeignete Präzisierungen des Utilitätsprinzips, wie sie etwa von *Harsanyi* (1977) oder — im deutschsprachigen Raum — von *Trapp* (1988) vorgeschlagen worden sind.

In diesen Grundprinzipien treten (wie zuvor dargestellt) Terme auf, deren Bedeutung im Basiselement des Netzes unterbestimmt bleibt. Im Fall des Utilitätsprinzips gilt das etwa für den Individuen- oder Nutzenbegriff. Die Bedeutung dieser Begriffe ist — wie im vorigen Abschnitt angedeutet — von den utilitaristischen Autoren in sehr unterschiedlicher Weise bestimmt worden und kann deswegen nicht bereits im Basiselement des Netzes festgehalten werden.

Eine nähere Bedeutungsfestlegung vollzieht sich in spezialisierten Theorie-Elementen. So gabelt sich das Netz der utilitaristischen Ethik am Basiselement in zwei Äste. Auf dem einen Ast liegen diejenigen spezialisierten Theorie-Elemente, in denen der Nutzenbegriff im Rahmen einer Theorie des psychologisch / ethischen Hedonismus interpretiert wird — wie im ausgehenden 18. und in der ersten Hälfte des 19. Jahrhunderts allgemein üblich. Auf dem anderen Ast liegen dagegen solche Theorie-Elemente, bei denen die Interpretation des Nutzenbegriffs auf einem geeignet präzisierten Präferenzbegriff aufbaut. Auf diesem Ast liegen insbesondere nahezu alle Varianten des Utilitarismus, die im zwanzigsten Jahrhundert entwickelt wurden. (Eine Ausnahme stellt A. C. Pigou dar, der in seinem wohlfahrtsökonomischen Fragestellungen gewidmeten Hauptwerk die hedonistische Tradition des Utilitarismus wiederbelebt hat; vgl. *Pigou* [4]1932). Der Vorgang der Spezialisierung kann iteriert werden. So kann etwa auf dem hedonistischen Ast weiter zwischen quantitativ-hedonistischen Varianten (Beispiel: Bentham) und qualitativ-hedonistischen Varianten (Beispiel: Mill) unterschieden werden. Dasselbe gilt für den präferenzutilitaristischen Ast: Hier kann weiter differenziert werden je nach Wahl des Präferenzbegriffs. Ebenso kann — wie in 6.1 dargestellt — zwischen verschiedenen Interpretationen des Individuenbegriffs unterschieden werden.

Wesentlich ist nicht nur, welche Theorie-Elemente im Netz utilitaristischer Ethiken vorkommen — aufschlußreich ist auch, welche Theorie-Elemente gerade nicht auftreten können. Daß manche Theorie-Elemente nicht auftreten können, hängt damit zusammen, daß die Interpretationen des Individuenbegriffs einerseits sowie des Nutzenbegriffs andererseits nicht unabhängig voneinander gewählt werden können. Sie sind vielmehr korreliert: Eine gegebene Interpretation des Individuenbegriffs erfordert eine Interpretation des Nutzenbegriffs, die auf alle Individuen dieses Grundbereichs auch tatsächlich anwendbar ist. Umgekehrt legt jede Interpretation des Nutzenbegriffs eine maximale Klasse berücksichtigungsfähiger Individuen fest. Die Korrelation von Nutzen- und Individuenbegriff bzw. der zugehörigen peripheren Theorieteile schränkt den Interpretationsspielraum

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/



für diese beiden Grundbegriffe der utilitaristischen Ethik deutlich ein. Damit senkt sie zugleich deren Immunisierungspotential erheblich.

Unter Berücksichtigung dieser Korrelationen können im Raster eines Theorien-Netzes durch iterierte Spezialisierung sukzessive feinere Varianten der utilitaristischen Position erfaßt und präzisiert werden. Damit können insbesondere unter philosophiehistorischen Gesichtspunkten relevante Positionen genauer und adäquater lokalisiert werden. Dieser Sachverhalt hängt wesentlich damit zusammen, daß das Konzept der *Theorien-Netze* die Einbettung ethischer Theorien in den Hintergrund des erfahrungswissenschaftlichen Wissens der betreffenden Epoche außerordentlich erleichtert: Integraler Bestandteil der Theorie-Elemente sind sog. *Links,* die es ermöglichen, Verbindungen zwischen ethischen Theorien und anderen — insbesondere empirischen — Vortheorien detailliert zu erfassen. Im Fall der utilitaristischen Ethik sind dies vor allem ökonomische und (motivations-)

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

psychologische Theorien, von deren Resultaten die utilitaristische Ethik in erheblichem Maße Gebrauch gemacht hat (vgl. *Gähde* [1993a]). Insbesondere wird es möglich nachzuzeichnen, wie Fortschritte und Modifikationen in den erfahrungswissenschaftlichen Vortheorien grundlegende Veränderungen im terminologischen Apparat utilitaristischer Ethiken induziert haben (vgl. dazu Abschnitt 6.2). Weiterhin ermöglichen es *Links,* die Beziehungen zwischen verschiedenen Varianten der utilitaristischen Position — etwa approximative Einbettungsbeziehungen — präziser zu erfassen. Es liegt nahe, ein Beispiel für eine derartige approximative Einbettungsbeziehung in dem fließenden Übergang zwischen regelutilitaristischen Positionen mit sukzessive enger gefaßten Handlungstypen und handlungsutilitaristischen Konzeptionen zu vermuten.

Die Vorteile einer umfassenden, präzisen Rekonstruktion des Komplexes utilitaristischer Theorien im Raster der *Theorien-Netze* zeigen sich vor allem dann, wenn man Konflikte zwischen diesen Theorien und basalen moralischen Intuitionen analysieren will. Jeder derartige Konflikt stellt zunächst einen Angriff auf die utilitaristische Ethik dar und kann eine Bedeutungsverschiebung der Grundtermini induzieren. Die Rekonstruktion der utilitaristischen Position als Theorien-Netz bietet erstens die Möglichkeit, die Zielpunkte derartiger Angriffe genauer zu lokalisieren. Vermeintlich allgemeingültige Gegenargumente gegen „den Utilitarismus" stellen sich häufig als Angriffe heraus, die sich nur auf spezielle Varianten richten und wenig über generelle Schwächen oder Vorzüge dieses Programms aussagen. Man kann so zeigen, welche Spielarten des Utilitarismus von einem Konflikt mit der Intuition betroffen sind und welche nicht.

Zweitens kann man mit Hilfe dieses Konzepts sowohl historisch *de facto* aufgetretene als auch theoretisch mögliche Verteidigungsstrategien differenziert erfassen und die durch sie bewirkten Bedeutungsverschiebungen etwa des Individuen- oder Nutzenbegriffs nachzeichnen. Dazu gehören insbesondere Fälle, in denen der Anwendungsbereich einer normativen Theorie erweitert und so eine Neubestimmung bzw. Präzisierung der Schlüsselbegriffe (etwa des Begriffs der Individuen, die als Objekte der moralischen Fürsorge zu gelten haben) erforderlich wird.

Drittens kann schließlich mit Hilfe dieses Konzepts die Frage angegangen werden, wo die Grenzen der Immunisierbarkeit des utilitaristischen Programms liegen. Der Grundgedanke ist dabei folgender: Von deskriptiven Theorien her ist bekannt, daß Korrekturversuche im allgemeinen bei den hochspezialisierten Theorie-Elementen eines Netzes beginnen und sich erst dann, wenn diese Modifikationen erfolglos bleiben, sukzessive auf weniger spezialisierte, fundamentalere Theorie-Elemente zubewegen. Es liegt nahe, die Grenze der Immunisierbarkeit da anzusetzen, wo anti-intuitive Konsequenzen nur um den Preis der Modifikation des Basiselementes des betreffenden Netzes vermieden werden können (das die *Identität* des betreffenden Theorienkomplexes bestimmt). Untersuchungen an empirisch-deskriptiven Theorien zeigen allerdings, daß für die Ermittlung der

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Grenzen des Immunisierungspotentials außerordentlich präzise Formulierungen des Basiselements vorliegen müssen. Es liegt nahe, entsprechende Untersuchungen versuchsweise auf den Fall normativ-ethischer Konzeptionen zu übertragen.

9. Metaethische Voraussetzungen

Es sei abschließend betont, daß die Anwendung wissenschaftstheoretischer Konzepte auf ethische Theorien sinnvoll ist unabhängig davon, ob man die normative Kraft basaler moralischer Intuitionen anerkennt oder nicht. Für den Intuitionisten, der grundlegenden Moralurteilen zumindest eine begrenzte Korrektivkraft in bezug auf die ethische Theorienbildung zubilligt, bietet das Raster der *Theorien-Netze* die Möglichkeit, seine moralischen Intuitionen zu systematisieren und damit den Aufbau seiner normativ-ethischen Theorie voranzutreiben. Für den Anti-Intuitionisten, der diese Korrektivkraft bestreitet (prominentes Beispiel: R. Hare; vgl. etwa *Hare* [1988]), liefern Theorien-Netze ein Modell, mit dem die faktische Entwicklung ethischer Positionen präziser nachgezeichnet und damit Ethikgeschichte transparenter gemacht werden kann.

Literatur

Bain, A.: James Mill. A Biography, London 1882, Nachdruck New York 1966.

Balzer, W. / *Moulines,* C. U. / *Sneed,* J.: An Architectonic for Science. The Structuralist Program, Dordrecht 1987.

Bentham, J.: A Fragment on Government (1776), Westport 1980.

— An Introduction to the Principles of Morals and Legislation (1789), hg. von J. H. Burns und H. L. A. Hart, London 1970.

— The Works of Jeremy Bentham, hg. von John Bowring, Edinburgh 1838-43.

— Jeremy Benthams Economic Writings, hg. von W. Stark, 3 Bde. London 1952-54.

Carnap, R.: Philosophical Foundations of Physics, New York 1966, dt. Einführung in die Philosophie der Naturwissenschaft, Frankfurt a.M. 1986.

Cumberland, R.: De Legibus Naturae (1672), englische Übersetzung von John Tower, Dublin 1750.

Duhem, P.: Ziel und Struktur der physikalischen Theorien, Leipzig 1908.

Dworkin, R.: Taking Rights Seriously, London 1977.

Eckensberger, L. H. / *Gähde,* U. (Hg.): Ethische Norm und empirische Hypothese, Frankfurt a.M. 1993.

Fechner, G. T.: Elemente der Psychophysik, Leipzig 1860.

Frankena, W. K.: „The Naturalistic Fallacy", Mind 48 (1939), 464-477.

Frey, R. G. (Hg.): Utility and Rights, Oxford 1985.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Gähde, U. / *Stegmüller*, W.: „An Argument in Favour of the Duhem-Quine-Thesis: From the Structuralist Point of View", in: Hahn, L. E. / Schilpp, A.: The Philosophy of W. V. Quine, The Library of Living Philosophers, vol. 15, La Salle, Il. 1986.

— „On Innertheoretical Conditions For Theoretical Terms", Erkenntnis 32 (1990), 215-233,

Gähde, U. / *Schrader*, H. (Hg.): Der klassische Utilitarismus: Einflüsse — Entwicklungen — Folgen, Berlin 1992.

— „Zum Wandel des Nutzenbegriffs im klassischen Utilitarismus", in: Gähde / Schrader (1992) (zitiert als *Gähde* 1992a).

— „Empirische und normative Aspekte der klassischen utilitaristischen Ethik", in: Ekkensberger / Gähde (1993), 63-91 (zitiert als *Gähde* 1993b).

Gossen, H. H.: Entwicklung der Gesetze des menschlichen Verkehrs und der daraus fließenden Regeln für menschliches Handeln (1854), 3. Aufl. Berlin 1927.

Hackmann, J.: „Allgemeine Akzeptierbarkeit als Auswahlkriterium für Werturteile", Jahrbuch für Sozialwissenschaft, Band 27 (1976).

Hare, R.: Moral Thinking. Its Levels, Method, and Point, Oxford 1988.

Harsanyi, J. C.: Rational Behavior and Bargaining Equilibrium in Games and Social Situations, Cambridge 1977.

Hegselmann, R.: „Solidarität in einer egoistischen Welt — eine Simulation", in: J. Nida-Rümelin / U. Wessels (Hg.), Praktische Rationalität, Berlin 1994.

Helvétius, C. A.: Œuvres Complètes, I-VI, hg. von Y. Belaval, Nachdruck der Ausgabe Paris 1795, Hildesheim, 1967-1969.

Hollis, M. / *Vossenkuhl*, W. (Hg.): Moralische Entscheidung und rationale Wahl, München 1992.

Lakatos, I. / *Musgrave*, A. (Hg.): Kritik und Erkenntnisfortschritt, Braunschweig 1974.

Maupertuis, P. L. M. de: Oeuvres, Bd. I-IV, Reprint der Ausgaben Lyon (1768) und Berlin (1758), Hildesheim / New York 1965-74.

McReynolds, P.: „The Motivational Psychology of Jeremy Bentham", Teil I: „Background and General Approach", Teil II: „Efforts Toward Quantification and Classification", Journal of the History of the Behavioural Sciences 4 (1968), 230-244 und 349-364.

Mill, J. St.: The Collected Works of John Stuart Mill, allgemeiner Herausgeber J. M. Robson, London / Toronto 1963-91.

Paley, W.: The Principles of Moral and Political Philosophy (1785), dt. Grundsätze der Moral und Politik, aus dem Engl. übers. von C. Garve, Nachdruck Hildesheim / Zürich 1986.

Pigou, A. C.: The Economics of Welfare, 4. Aufl. London 1932.

Pollard, S.: „Der klassische Utilitarismus: Einflüsse, Entwicklungen, Folgen", in: Gähde / Schrader (1992).

Popper, K. R.: Logik der Forschung, 1. Aufl. Wien 1934 (mit der Jahresangabe 1935), 8. verbesserte und vermehrte Aufl. Tübingen 1984.

Rawls, J.: A Theory of Justice, Cambridge, Mass. 1971; dt. Eine Theorie der Gerechtigkeit, Frankfurt a.M. 1975.

Sen, A. K.: Collective Choice and Social Welfare, San Francisco 1970.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

— „Utilitarianism and Welfarism", The Journal of Philosophy, Vol. LXXVI, No. 9 (1979).

Sidgwick, H.: The Methods of Ethics, 1. Auflage 1874, 7. Auflage (mit einem Vorwort von John Rawls), Indianapolis-Cambridge 1981.

Sikora, R. I. / *Barry,* B. (Hg.): Obligations to Future Generations, Philadelphia 1978.

Smart, J. J. C. / *Williams,* B.: Utilitarianism — for and against, Cambridge 1973.

Spohn, W.: „Wie kann die Theorie der Rationalität empirisch und normativ zugleich sein?", in: Eckensberger / Gähde (1993).

Stegmüller, W.: Theorie und Erfahrung, dritter Teilband, Berlin-Heidelberg / New York / Tokyo 1986.

Trapp, R.: Nicht-klassischer Utilitarismus. Eine Theorie der Gerechtigkeit, Frankfurt a.M. 1988.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Der Schleier der Unwissenheit und die Idee eines übergreifenden Konsenses

Von *Wilfried Hinsch,* Münster

1.

In *Political Liberalism* (PL)[1] entwickelt John Rawls die normativen Grundlagen einer vertragstheoretischen Konzeption politischer Legitimität. Ausgehend von seiner Theorie der Gerechtigkeit als Fairneß geht es ihm darum, zu zeigen, daß eine gerechte und legitime politische Ordnung auch unter den für moderne Demokratien typischen Bedingungen eines moralischen und religiösen Pluralismus möglich ist. Das von Rawls vorgestellte Ideal politischer Ordnung ist in seiner Konzeption einer wohlgeordneten Gesellschaft verkörpert. Eine wohlgeordnete Gesellschaft ist ein Generationen übergreifendes System sozialer und politischer Kooperation, das von öffentlich anerkannten Gerechtigkeitsgrundsätzen wirksam reguliert wird. Im einzelnen müssen folgende Bedingungen erfüllt sein: (1) Alle gesellschaftlichen Institutionen und Regeln entsprechen den von allen Bürgern aus guten Gründen anerkannten Gerechtigkeitsgrundsätzen. (2) Alle handeln gesetzestreu und in Übereinstimmung mit der öffentlichen Gerechtigkeitsauffassung. (3) Alle wissen, daß die Bedingungen (1) und (2) erfüllt sind, und auch dies ist allgemein bekannt. (4) Es besteht Einigkeit über die Begründung der obersten Gerechtigkeitsgrundsätze (PL: 35 ff., 66 f.).

Der Rawls'schen Konzeption politischer Legitimität liegt ein Begriff der *öffentlichen Rechtfertigung* von Normen zugrunde. Grundsätze politischer Gerechtigkeit müssen auf der Basis von allen geteilter normativer Überzeugungen die öffentliche Zustimmung derjenigen finden, für die sie als oberste regulative Prinzipien der wichtigsten gesellschaftlichen Institutionen Geltung beanspruchen. Eine legitime politische Ordnung zeichnet sich demnach dadurch aus, daß die für sie konstitutiven Regeln und Grundsätze vor allen Bürgern, allen gegenüber mit für alle gleichermaßen einsichtigen Gründen gerechtfertigt werden können (PL: 66 f.). Dies entspricht dem vertragstheoretischen Grundgedanken der Gerechtigkeit als Fairneß: gerecht ist, worauf sich freie und gleiche Bürger in einer fairen Übereinkunft einigen würden.

---

[1] New York: Columbia University Press, 1993.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Forderung der Öffentlichkeit spielte schon für das vertragstheoretische Argument in *A Theory of Justice* (TJ)[2] eine wichtige Rolle. Sie wurde dort jedoch vor allem unter dem Gesichtspunkt der Stabilität einer wohlgeordneten Gesellschaft diskutiert.[3] Stabile politische und soziale Kooperation setzt voraus, daß die Beteiligten nicht nur *de facto* bestimmte Regeln anerkennen und befolgen, sondern auch voneinander wissen, daß sie dies tun. Das wird durch die Öffentlichkeitsbedingung sichergestellt.[4] Darüber hinaus geben Bürger durch ihre öffentliche Anerkennung fairer Grundsätze politischer Kooperation voreinander zu erkennen, daß sie einander als gleichberechtigte Partner betrachten, und auch darin liegt ein Beitrag zur Stabilisierung einer wohlgeordneten Gesellschaft.

In *Political Liberalism* gewinnt der Gedanke der öffentlichen Rechtfertigung nun eine zusätzliche Dimension dadurch, daß das ‚Faktum des Pluralismus' ausdrücklich als ein Grundproblem sowohl der Begründung allgemeiner politischer Grundsätze als auch der Legitimierung konkreter politischer Entscheidungen anerkannt wird. Es geht nicht mehr allein darum, daß nur öffentlich anerkannte Grundsätze politischer Gerechtigkeit zu einem stabilen System sozialer Kooperation führen. Die vorrangige Frage ist jetzt, unter welchen Bedingungen es in einer pluralistischen Gesellschaft grundsätzlich möglich erscheint, öffentlich begründbare Gerechtigkeitsgrundsätze allererst zu identifizieren.

2.

Die von Rawls vertretene vertragstheoretische Auffassung politischer Gerechtigkeit führt, sobald das ‚Faktum des Pluralismus' als historische und normative Grundtatsache anerkannt wird, zum Postulat eines ‚übergreifenden Konsenses'. Keine der von den Bürgern moderner demokratischer Staaten vertretenen umfassenden moralischen oder religiösen Lehren findet allgemeine Zustimmung, und keine von ihnen kommt deshalb als Grundlage für eine Gerechtigkeitskonzeption in Betracht, die allen gegenüber gerechtfertigt werden könnte. Die öffentliche Rechtfertigung politischer Grundsätze kann gleichwohl nur dann gelingen, wenn die von den Bürgern bejahten umfassenden normativen Konzeptionen unangesehen ihrer sonstigen Divergenzen hinreichend viele Übereinstimmungen aufweisen. Denn zur Ausarbeitung einer vertragstheoretischen Gerechtigkeitskonzeption können nur solche Voraussetzungen in Anspruch genommen werden, die von einem ‚übergreifenden Konsens' getragen werden oder über die im Prinzip ein solcher Konsens durch rationale Argumentation herbeigeführt werden kann. In den Worten von John Rawls: „this political conception needs to be such that there is some hope of its gaining the support of an overlapping consensus, that

---

2 Cambridge, Mass.: Harvard University Press 1971.

3 TJ: 133, 177-182, 582, vgl. auch 454, 547 f., 570.

4 Vgl. oben das dritte definierende Merkmal einer wohlgeordneten Gesellschaft.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

is, a consensus in which it is affirmed by the opposing religious, philosophical or moral doctrines likely to thrive over generations in a more or less just constitutional democracy."[5]

Die Bedeutung von Zustimmung und Konsens für Gerechtigkeitsgrundsätze ist kaum zu übersehen, wenn man deren gesellschaftliche Aufgabe bedenkt, ein System von Institutionen zu regulieren, das faire und fruchtbare soziale Kooperation ermöglicht. Jede Form dauerhafter politischer Ordnung setzt ein hohes Maß freiwilliger Kooperation und unerzwungener Regelbefolgung voraus. Die zwangsweise Durchsetzung von Gesetzen und politischen Maßnahmen durch den Staat bringt hohe Opportunitätskosten mit sich und stößt schnell an ihre Grenzen. Kein Regime würde lange bestehen, wenn es sich nicht darauf verlassen könnte, daß hinreichend viele Bürger die ihm zugrundeliegenden Prinzipien zumindest stillschweigend gutheißen und deshalb seinen Geboten und Forderungen willig Folge leisten.

Die stabilitätsbezogene, an Hobbes orientierte Konzeption politischen Konsenses betrachtet die Zustimmung der Bürger deshalb zu Recht als eine empirisch-kausale Bedingung dauerhafter Ordnung. Worauf es unter diesem Gesichtspunkt ankommt, ist nicht, daß alle Bürger das Regime bejahen und unterstützen, sondern lediglich, daß hinreichend viele es tun. Einstimmigkeit ist unnötig, da stabile politische Verhältnisse normalerweise auch ohne sie erreicht werden können. Sobald die Anzahl der ‚loyalen' Bürger eine kritische Schwelle überschreitet, entstehen für alle Gesellschaftsmitglieder unabhängig von ihrer Einstellung gegenüber dem Regime Bedingungen, die offenen Widerstand als riskant und aussichtslos erscheinen lassen.

Eine andere Perspektive ergibt sich, wenn man von Kant ausgeht und wie er die Forderung, das ganze Volk müsse seine ‚Einstimmung' geben können,[6] aus dem Gedanken der Autonomie des Individuums herleitet. Die grundlegende Idee dabei ist, daß Bürger als freie und gleiche moralische Personen einen unabweisbaren Anspruch haben, nur solchen Normen unterworfen zu werden, denen sie auf der Grundlage ihrer wohlerwogenen Interessen und Überzeugungen zustimmen können. Die Möglichkeit eines begründeten Konsenses wird dabei zu einem konstitutiven Element politischer Gerechtigkeit, und es reicht nicht, daß die Anzahl der zustimmenden Bürger eine kritische Schwelle überschreitet, sondern *alle* Bürger müssen zustimmen können.

Für die empirische Stabilität einer institutionellen Ordnung sind die faktischen Einstellungen und Überzeugungen der Beteiligten entscheidend. Deren Rationalität spielt allenfalls indirekt eine Rolle. Wenn Bürger ein Regime unterstützen,

---

[5] „The Idea of an Overlapping Consensus" (IOC), (Oxford Journal of Legal Studies 7, (1987), S. 1-25), S. 1.

[6] So *Kant* in der Schrift über das Verhältnis von Theorie und Praxis im Staatsrecht, Akad. Ausg. Bd. 8, S. 297.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

weil sie glauben, es sei ihre Pflicht oder in ihrem eigenen Interesse, dies zu tun, so tragen sie damit zur Stabilität des Regimes bei, egal, ob ihre subjektiven Überzeugungen konsistent oder inkonsistent, begründet oder unbegründet sind. Der kantischen Konzeption von Zustimmung und Konsens dagegen geht es nicht um faktische Anerkennung, sondern um die moralische Rechtfertigung von politischen Grundsätzen und Institutionen. Als gerechtfertigt gelten diese aber erst dann, wenn alle Bürger ausschlaggebende Gründe haben, sie als verbindlich für sich und andere anzuerkennen, und nicht schon, wenn alle *glauben,* solche Gründe zu haben.

Meine stark vereinfachte Unterscheidung zwischen einer hobbes'schen und einer kantischen Auffassung über die Notwendigkeit eines Konsenses soll nicht zur Klassifikation politischer Theorien dienen. Ich gehe davon aus, daß jede adäquate Theorie politischer Gerechtigkeit hobbes'sche und kantische Elemente enthält. Worum es geht, ist, deutlich zu machen, zwischen welchen Polen sich eine theoretische Bestimmung der Rolle von Zustimmung und Konsens in der politischen Philosophie bewegt.

Nachdem Rawls 1987 in „The Idea of an Overlapping Consensus"[7] seine Konzeption eines übergreifenden Konsenses zuerst ausführlich vorgestellt hatte, wurde ihm von Kritikern vorgeworfen, die Begründung von Gerechtigkeitsgrundsätzen in unzulässiger Weise mit pragmatisch-politischen Erwägungen zu durchsetzen.[8] Tatsächlich betont Rawls in seinen neueren Arbeiten die ‚praktische Aufgabe' der politischen Philosophie.[9] Diese sieht er darin, eine Gerechtigkeitskonzeption auszuarbeiten, die „von Bürgern als gemeinsame Basis für eine begründete, informierte und freiwillige politische Übereinkunft geteilt werden kann" (PL: 9), und die deshalb in der Lage ist, auf der Basis eines übergreifenden Konsenses die soziale Einheit und Stabilität pluralistischer Gesellschaften zu gewährleisten. So entstand der Eindruck, als ob Rawls seine Gerechtigkeitsgrundsätze als das Ergebnis eines Kompromisses zwischen den in einer Gesellschaft vorherrschenden moralischen und religiösen Auffassungen betrachte. Dann aber wäre der Inhalt politischer Gerechtigkeit abhängig von historisch kontingenten Übereinstimmungen zwischen divergierenden und zum Teil einander ausschließenden Lehren und Anschauungen. Und dies ließe in der Tat befürchten, die politische Philosophie könnte auf die falsche Weise politisch werden. Es liefe darauf hinaus, um der sozialen Einheit und Stabilität willen die Forderungen der Gerechtigkeit den Ideologien der vorherrschenden gesellschaftlichen Gruppen anzupassen.

---

7 Vgl. Fn. 5.

8 Vgl. etwa *Joseph Raz* in: „Facing Diversity. The Case of Epistemic Abstinence", (Philosophy and Public Affairs 19, [1990], S. 3-46), der stellvertretend für viele die Frage stellt, „. . . Has Rawls become a Politician?" (S. 10).

9 PL: 10, vgl. „Kantian Constructivism in Moral Theory" (KC), (Journal of Philosophy 77, [1980], S. 515-572), S. 518 f., 538, 541, 543; IOC: 13n, 15.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Um diesem Mißverständnis entgegenzutreten, werden in *Political Liberalism* zwei Darstellungsstufen unterschieden. Auf der ersten Stufe soll ‚Gerechtigkeit als Fairneß' ohne Einbeziehung des Gedankens eines übergreifenden Konsenses als eine, wie Rawls sagt, „freistehende politische Konzeption" vorgestellt und begründet werden. Erst im Anschluß daran ist dann vom übergreifenden Konsens die Rede, wenn es darum geht, zu erörtern, unter welchen Bedingungen eine wohlgeordnete pluralistische Gesellschaft dauerhaft bestehen kann (PL: 64 f. u. 133 f.). Die Unterscheidung einander nachgeordneter Darstellungsstufen entspricht der tatsächlichen Entwicklung der Rawls'schen Theorie in den letzten zehn Jahren und ist geeignet, den angeführten Einwand zu entkräften. Sie gibt aber keinen angemessenen Begriff von der Bedeutung eines übergreifenden Konsenses für die Formulierung einer vertragstheoretischen Gerechtigkeitsauffassung. Im Gegenteil scheint sie die Annahme zu bestätigen, ein übergreifender Konsens sei zunächst und vor allem unter Gesichtspunkten politischer Stabilität erforderlich.[10] Doch genau dies ist falsch. Wie sich im folgenden zeigen wird, sprechen prinzipielle Überlegungen, die an den Begriff der öffentlichen Rechtfertigung und an Rawls' Herleitung seiner beiden Gerechtigkeitsgrundsätze mit Hilfe der Original Position anknüpfen, für eine kantische Interpretation des Postulats eines übergreifenden Konsenses.

3.

3.1 Um meine These zu begründen, beginne ich mit einer kursorischen Darstellung des Rawls'schen Argumentationsmodells der Original Position. Dabei orientiere ich mich an dem, was Rawls zuerst 1980 in den Dewey-Lectures über die normativen Voraussetzungen seines Modells gesagt hat (KC: 516-528, vgl. PL: 18-35 u. 72-77).

Die Original Position beschreibt eine hypothetische Vertragssituation. Die Bürger einer Gesellschaft (oder deren Vertreter) beraten hinter einem ‚Veil of Ignorance' darüber, nach welchen obersten Prinzipien sie die institutionelle Grundstruktur ihrer Gesellschaft einrichten wollen. Zur Auswahl stehen ihnen die aus der Tradition bekannten Grundsätze politischer und sozialer Gerechtigkeit, die sie, falls nötig, abwandeln können. Wegen des ‚Schleiers der Unwissenheit' wissen die Parteien in der Original Position weder, welche persönlichen Eigenschaften, Fähigkeiten und Interessen sie haben, noch kennen sie ihre besonderen moralischen Überzeugungen und Wertvorstellungen; darüber hinaus ist ihnen die Gesellschaft, in der sie leben, unbekannt, und auch ihre konkrete soziale Stellung kennen sie nicht. Sie wissen aber, daß sie freie und gleiche Bürger sind und über zwei moralische Vermögen verfügen, deren Entwicklung und Ausübung für sie von überragender Bedeutung ist: den Gerechtigkeitssinn und die Fähigkeit

[10] Vgl. PL: xvf., 39, aber auch die Klarstellung gegenüber Baier auf S. 149.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

zur Ausbildung, kritischen Überprüfung und Verfolgung einer Konzeption des Guten.

Der Begriff des Bürgers als freier und gleicher moralischer Person ist für Rawls' Konzeption grundlegend. Er gehört zu den ‚fundamental intuitive ideas', von denen Rawls annimmt, daß sie zumindest implizit allgemein anerkannt werden, und von denen ausgehend es ihm möglich erscheint, eine Gerechtigkeitskonzeption zu formulieren, die ebenfalls mit allgemeiner Zustimmung rechnen kann (KC: 516-518, PL: 18 f.). Moralische Personen zeichnen sich durch drei definierende Merkmale aus: Sie verfügen (1) über einen Gerechtigkeitssinn, (2) über das Vermögen zu einer Konzeption des Guten und (3) über eine mehr oder weniger umfassende und mehr oder weniger artikulierte besondere Konzeption des Guten. Diesen drei Merkmalen korrespondieren drei übergeordnete Interessen (‚higher-order interests') moralischer Personen. Sie bilden die motivationale Basis für die Auswahl von Gerechtigkeitsgrundsätzen durch die Parteien in der Original Position. Jedem der beiden moralischen Vermögen ist ein Interesse an seiner adäquaten Ausbildung und Ausübung zugeordnet. Der bestimmten, aber vom Standpunkt der Original Position aus unbekannten Konzeption des Guten jedes Bürgers entspricht ein Interesse daran, den eigenen Lebensentwurf, worin immer er bestehen mag, gegen Übergriffe zu schützen und so gut es geht zu verwirklichen (KC: 525, PL: 19, 73 f.).

Da den Parteien hinter dem ‚Schleier der Unwissenheit' alle Informationen über ihre persönlichen und sozialen Lebensumstände fehlen, orientieren sie sich bei der Auswahl von Gerechtigkeitsgrundsätzen ausschließlich an den drei genannten übergeordneten Interessen. So wählen sie Grundsätze, von denen sie aus der Perspektive der Original Position annehmen, daß sie die für die Ausbildung und Ausübung ihrer beiden moralischen Vermögen und für die Förderung ihrer jeweiligen Konzeption des Guten notwendigen sozialen Bedingungen gewährleisten.

Rawls beschreibt diese Bedingungen als ‚Grundgüter', an denen alle Bürger, als moralische Personen, ein rationales Interesse haben, weil sie für die Wahrnehmung ihrer übergeordneten Interessen unverzichtbar sind. „These goods, we say, are things citizens need as free and equal persons . . .".[11] Welche Güter in diesem Sinne als ‚Grundgüter' gelten, wird von Rawls in Form einer Liste festgelegt, die den Beratungen der Parteien in der Original Position zugrundeliegt. Sie umfaßt: (1) Bürgerliche und politische Grundrechte, (2) Freizügigkeit und freie Berufswahl, (3) Befugnisse und Vorrechte von Ämtern und Positionen, (4) Einkommen und Besitz, verstanden als allgemein dienliche Mittel zur Verfolgung

[11] KC: 525 ff., PL: 75 f. u. 176-190. Zur Rawls'schen Konzeption der Grundgüter in ihrer gegenüber *Eine Theorie der Gerechtigkeit* revidierten Form vgl. die ausführliche Diskussion in meiner Einleitung zu: *John Rawls,* Die Idee des politischen Liberalismus. Aufsätze 1978-1989, (Frankfurt a.M.: Suhrkamp Verlag, 1992), S. 36-44.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

einer Vielzahl möglicher Lebenspläne, (5) die sozialen Grundlagen der Selbstachtung.

Wegen der überragenden Bedeutung einer adäquaten Ausbildung der beiden Vermögen und angemessener Bedingungen für ihre Ausübung beurteilen die Parteien die ihnen vorliegenden Gerechtigkeitsgrundsätze aus der Perspektive der am wenigsten begünstigten Gesellschaftsmitglieder — d.h. derjenigen, die de facto den kleinsten Anteil von Grundgütern für sich reklamieren können — und wählen die Grundsätze in Übereinstimmung mit der sog. Maximin-Regel der rationalen Entscheidungstheorie, das heißt, sie bemühen sich, für die am wenigsten Begünstigten den für sie größtmöglichen Anteil an Grundgütern zu sichern. Unter diesen Bedingungen, so Rawls, würden sich die Parteien auf die von ihm vorgeschlagenen beiden Gerechtigkeitsgrundsätze einigen: Der erste Grundsatz fordert ein adäquates System gleicher politischer und bürgerlicher Freiheiten für alle, der zweite faire Chancengleichheit und die größtmögliche Zuteilung an Grundgütern für die am wenigsten begünstigten Gesellschaftsmitglieder (vgl. PL: 5).

Die Konstruktion der Original Position, wie ich sie beschrieben habe, läßt leicht erkennen, auf welchen hauptsächlichen Vorausetzungen die Herleitung der beiden Grundsätze beruht: es sind erstens der Begriff des Bürgers als freier und gleicher moralischer Person und zweitens der dichte ‚Schleier der Unwissenheit', der die vertragschließenden Parteien nicht nur von Informationen über ihre persönlichen Präferenzen und Interessen abschirmt, sondern auch jede Berufung auf umfassende moralische und religiöse Konzeptionen des Guten ausschließt.

Erklärungsbedürftig erscheint zunächst, welche Relevanz die Überlegungen und Entscheidungen der Parteien in der Original Position — die Überlegungen also von künstlich konzipierten Gedankenwesen in einer nicht minder künstlichen Situation — für Personen haben, die nicht hinter einem ‚Schleier der Unwissenheit' leben, die also ihre persönlichen Präferenzen, Interessen und Überzeugungen kennen. Doch die Antwort ist offensichtlich: Wenn das Rawls'sche Konstrukt argumentative Kraft hat, dann deshalb, weil wir die in seinen Elementen verkörperten Annahmen anerkennen oder nach gebührender Überlegung anerkennen würden.

3.2 Das Vermögen moralischer Personen zu einer Konzeption des Guten mit den korrespondierenden übergeordneten Interessen erscheint als Voraussetzung zunächst unproblematisch. Es wird, soweit ich sehen kann, allgemein anerkannt, daß die Fähigkeit, Auffassungen darüber auszubilden, was gut und schlecht ist, und ihnen in mehr oder weniger rationaler Weise nachzuleben, eine notwendige Bedingung moralischer Personalität ist. Problematischer erscheint die Annahme eines Gerechtigkeitssinnes. Sie wird von vielen im Namen der methodischen Konzeption des rationalen Egoismus in Frage gestellt und verdient eine ausführliche Erörterung. Ich begnüge mich jedoch mit einem kurzen Kommentar.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Ein rationaler Egoist ist jemand, für den eine Handlung genau dann begründet ist, wenn sie in seinem wohlverstandenen Eigeninteresse liegt. Im paradigmatischen Grenzfall nimmt ein rationaler Egoist niemals persönliche Nachteile in Kauf, es sei denn um längerfristiger Vorteile willen. Dies schließt nicht aus, daß die Existenz eines Systems sozialer Regeln, da sie im wohlverstandenen Eigeninteresse aller Gesellschaftsmitglieder liegt, auch einem rationalen Egoisten gegenüber begründet werden kann. Er hat zumindest ein Interesse daran, daß seine Mitbürger ihm gegenüber bestimmte Verhaltensbeschränkungen einhalten; denn nur unter dieser Voraussetzung darf er hoffen, seine eigenen Lebenspläne, worin immer sie bestehen mögen, verwirklichen zu können. Wir dürfen deshalb erwarten, daß ein rationaler Egoist erstens die Einrichtung eines sanktionsbewehrten Systems sozialer Regeln befürworten und unterstützen wird, und zweitens, daß er sich wegen der Opportunitätskosten regelwidrigen Verhaltens normalerweise selbst regelkonform verhalten wird. In diesem Sinne können wir davon sprechen, daß er die Geltung sozialer Regeln anerkenne.[12]

Entscheidend ist, daß der rationale Egoist Regeln ausschließlich unter instrumentellen Gesichtspunkten in seinem Handeln berücksichtigt, nämlich als notwendige empirische Rahmenbedingungen für die optimale Verwirklichung seiner Lebenspläne. Er erkennt sie in dem Sinne an, in dem wir auch davon sprechen, daß jemand die Tatsachen anerkenne: er nimmt sie zur Kenntnis, begrüßt gegebenenfalls ihr Bestehen und berücksichtigt sie in möglichst umsichtiger Weise bei der Verfolgung seiner Pläne. Er macht sie aber nicht zum Maßstab seines Handelns und wird sich nur insoweit an sie halten, wie dies seinen längerfristigen Interessen dient. Wann immer sich ihm die Möglichkeit bietet, durch gezielte Regelverletzungen Vorteile zu erzielen, wird er dies tun.

Man beachte, daß es einem rationalen Egoisten aus logischen Gründen unmöglich ist, Regeln in dem weiteren Sinne anzuerkennen, den wir mit moralischen Geboten gewöhnlich verbinden. Dazu würde gehören, daß das Gebot als ein *intrinsischer Grund* für regelkonformes Verhalten und damit als *Maßstab* des eigenen Verhaltens betrachtet wird. Und dies setzt voraus, davon überzeugt zu sein, daß man der Regel auch dann folgen soll, wenn dies gelegentlich mit unkompensierten persönlichen Nachteilen verbunden ist. Eine solche Überzeugung ist jedoch mit dem Begriff eines rationalen Egoisten logisch unvereinbar.

Einwände gegen die begründungstheoretische Annahme eines Gerechtigkeitssinnes, die auf der Konzeption des wohlverstandenen Eigeninteresses beruhen,

---

12 Dieser Punkt ist im Anschluß an Hobbes und Hume immer wieder hervorgehoben worden. Mit ihm beginnen alle Versuche, die elementaren Voraussetzungen sozialer Ordnung auf der Grundlage des rationalen Eigeninteresses der Beteiligten zu begründen. Vgl. *James Buchanan,* The Limits of Liberty, (The University of Chicago Press, 1975), *John Mackie,* Ethics. Inventing Right and Wrong, (Harmondsworth: Penguin Books, 1977), Kap. 5, und *David Gauthier,* Morals by Agreement, (Oxford: Clarendon Press 1986).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

scheitern demnach schon im Vorfeld daran, daß es unter rationalen Egoisten keine moralischen Rechtfertigungen im gewöhnlichen Sinne geben kann, weil diese Homunculi der neuzeitlichen Ökonomie niemals universalisierbaren Geboten aufrichtig und ernsthaft zustimmen.[13] Sobald wir uns auf der Ebene moralischen Argumentierens befinden, sind unsere Adressaten deshalb keine rationalen Egoisten, sondern moralische Personen im vollen Sinne, und dies findet seinen angemessenen Ausdruck in der Rawls'schen Konzeption der moralischen Person.

3.3 Die von den übergeordneten Interessen moralischer Personen ausgehenden Überlegungen der Parteien in der Original Position sind auch vor dem ‚Schleier der Unwissenheit', also ohne Informationsbeschränkungen, starke Gründe für die von Rawls favorisierten Gerechtigkeitsgrundsätze. Wenn die von Rawls benannten Grundgüter tatsächlich notwendige soziale Voraussetzungen der Entwicklung und Ausübung der beiden moralischen Vermögen sind, haben alle Gesellschaftsmitglieder angesichts stets ungewisser Zukunftserwartungen gute Gründe, *ceteris paribus* eine institutionelle Grundstruktur zu wünschen, die ihnen und ihren Kindern dauerhaft eine hinreichende Zuteilung an Grundgütern in allen gesellschaftlichen Positionen garantiert. Erst durch den dichten ‚Schleier der Unwissenheit' werden diese Gründe jedoch zu ausschlaggebenden Gründen, die alle entgegenstehenden Überlegungen überwiegen und definitiv zu den Rawls'schen Grundsätzen und dem mit ihnen festgeschriebenen Vorrang der Grundfreiheiten führen.

Man beachte in diesem Zusammenhang die logische Funktion des ‚Schleiers der Unwissenheit'. Eine schon Aristoteles vertraute Besonderheit praktischen Schließens besteht darin, daß für einen praktischen Schluß, dessen Konklusion eine präskriptive Komponente enthält, die Hinzufügung wahrer Prämissen ebenso verhängnisvoll sein kann, wie berechtigte Kritik an den Annahmen, auf denen der Schluß beruht. Je nachdem, welche Grundsätze, Werte und Tatsachen jemand in seinen Überlegungen berücksichtigt, können sich verschiedene praktische Folgerungen ergeben. Die Funktion des ‚Schleiers der Unwissenheit' besteht unter diesem Gesichtspunkt darin, die materialen Prämissen, aus denen die Gerechtigkeitsgrundsätze deduktiv gewonnen werden — also im wesentlichen die Konzeption der moralischen Person — davor zu schützen, durch Einbeziehung zusätzlicher Informationen ihrer ausschlaggebenden Kraft beraubt zu werden.[14] Jemand, der seine bevorzugte soziale Position kennt, hat, auch wenn er Rawls' Begriff der moralischen Person anerkennt, *ceteris paribus* gute Gründe, das ‚Differenzprinzip' abzulehnen. Und ein strenggläubiger Katholik könnte dem übergeordne-

---

[13] Unter einer moralischen Rechtfertigung ‚im gewöhnliche Sinne' verstehe ich ein praktisches Argument, dessen Konklusion ein universalisierbares Gebot enthält und das genau dann gelungen ist, wenn sein Adressat der Konklusion auf der Basis der angeführten Prämissen aufrichtig und ernsthaft zustimmt.

[14] Die Bedeutung des Veil of Ignorance für die ausschließliche Berücksichtigung der übergeordneten Interessen moralischer Personen wird von Rawls ausdrücklich hervorgehoben. Vgl. KC: 549 f.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ten Interesse, das er als moralische Person daran hat, seine Konzeption des Guten gegebenenfalls revidieren zu können, ein geringeres Gewicht beimessen als seinem konkreten Interesse daran, in Gemeinschaft mit Gleichgesinnten ein gottesfürchtiges Leben zu führen. Er würde sich dann nicht für den von Rawls geforderten Vorrang der Gewissensfreiheit entscheiden und gegebenenfalls die Einrichtung einer katholischen Staatskirche befürworten.

Für ein genaueres Verständnis des Veil of Ignorance ist es hilfreich, drei Stufen der durch ihn vorgenommenen Informationsbeschränkungen zu unterscheiden: Die *erste Stufe* betrifft den Ausschluß von Informationen über die persönlichen Präferenzen und konkreten Lebensumstände einzelner. Auf der *zweiten Stufe* werden außerdem alle Informationen über die allgemeinen gesellschaftlichen Lebensverhältnisse (ökonomischer Entwicklungsstand, Einkommensverteilung, Klassenstruktur) ausgeschlossen. Auf der *dritten Stufe* kommt es schließlich zum Ausschluß aller Informationen über die umfassenden moralischen, philosophischen und religiösen Überzeugungen der Gesellschaftsmitglieder.

Der von Rawls in Anspruch genommene *dichte* Veil of Ignorance realisiert alle drei Stufen, so daß sich die Frage nach der Berechtigung von Informationsbeschränkungen ebenfalls auf drei Ebenen stellt. Rawls hat den Veil of Ignorance als eine „natural condition that must have occurred to many" (TJ: 137 n., dt. Übers.: 159 f. Fn.) bezeichnet, und dies leuchtet ein, solange wir uns auf Stufe Eins beziehen, das heißt auf den Ausschluß von Informationen, die persönliche Präferenzen und Interessen der vertragschließenden Parteien betreffen. Ein ‚Schleier', der diese Informationen ausschließt, ist eine notwendige Voraussetzung der unparteiischen Auswahl von Gerechtigkeitsgrundsätzen und offenkundig gerechtfertigt, weil die persönlichen Präferenzen und besonderen Lebensumstände einzelner nicht einmal *prima facie* Gründe für die Wahl oberster Gerechtigkeitsgrundsätze abgeben.

Ähnliches gilt für die zweite Stufe, durch die Informationen über die allgemeinen gesellschaftlichen Lebensverhältnisse von den Beratungen in der Original Position ausgeschlossen werden. Dies ist notwendig, um zu verhindern, daß der Inhalt politischer Gerechtigkeit in unkontrollierbarer Weise von Faktoren beeinflußt wird, die selber erst das Resultat gerechter oder ungerechter historischer Entwicklungen und sozialer Strukturen sind. Die kritische Funktion oberster Gerechtigkeitsgrundsätze gegenüber der institutionellen Grundstruktur einer Gesellschaft schließt es aus, bei ihrer Identifikation von Faktoren auszugehen, die selber bereits das Resultat spezifischer institutioneller Festlegungen sind.[15] So

[15] Rawls hat diesen Gedanken mit Hilfe seiner Konzeption der ‚background justice' in „The Basic Structure as Subject" (1978) ausgeführt (vgl. PL: 265-275). Zu den Konsequenzen, die sich aus ihm für an Präferenzen orientierte Gerechtigkeitkeitstheorien ergeben, vgl. meine Kritik an Hares Theorie in: „Präferenzen im moralischen Denken" (1990), abgedr. in: *Fehige / Meggle* (Hrsg.), Zum moralischen Denken, (Frankfurt a.M.: Suhrkamp Verlag, 1994).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ist Rawls zuzustimmen, wenn er sagt: „we must find some point of view, removed from and not distorted by the particular features and circumstances of the all-encompassing background framework, from which a fair agreement . . . can be reached" (PL: 23).

Rawls versucht nun, den auf der dritten Stufe realisierten Ausschluß aller Informationen über umfassende moralische und religiöse Konzeptionen des Guten in analoger Weise zu rechtfertigen wie die nötigen Beschränkungen der ersten und zweiten Stufe.

> „Moreover, one of our considered convictions, I assume, is this: the fact that we occupy a particular social position is not a good reason for us to propose, or to expect others to accept, a conception of justice that favors those in this position. Similarly, the fact that we affirm a particular religious, philosophical, or moral comprehensive doctrine with its associated conceptions of the good is not a reason for us to propose, or to expect others to accept, a conception of justice that favors those of that persuasion. To model this conviction in the original position, the parties are not allowed to know the social position of those they represent, or the particular comprehensive doctrine of the person each represents" (PL: 24).

Doch dieser Übergang von der ersten zur dritten Stufe der Informationsbeschränkung beruht auf einem argumentativen Kurzschluß. Er setzt voraus, daß die Bestimmung von Gerechtigkeitsgrundsätzen unter Berücksichtigung nicht allgemein geteilter umfassender Konzeptionen des Guten zu derselben parteilichen Bevorzugung einzelner führt wie die Einbeziehung partikularer Interessen. Das ist jedoch nicht notwendigerweise der Fall. Eine von mir vertretene umfassende Konzeption des Guten ist nicht darum schon eine Konzeption, die mich bevorzugt, weil nur ich, und vielleicht einige Gleichgesinnte, sie vertreten. Dies wird von Rawls ausdrücklich anerkannt (PL: 15), und für die von ihm selbst immer wieder genannten paradigmatischen Beispiele für umfassende moralische und religiöse Lehren (Kants und Mills Liberalismus, Protestantismus und Katholizismus) gilt, daß sie einen aus ihrer Perspektive nicht unbegründeten Anspruch auf Unparteilichkeit erheben. Ob sich aus der Berücksichtigung umfassender moralischer und religiöser Überzeugungen eine ungerechtfertigte Bevorzugung einzelner ergibt oder nicht, kann deshalb nicht vorab, sondern nur auf der Basis einer bereits identifizierten politischen Gerechtigkeitskonzeption entschieden werden, einer Gerechtigkeitskonzeption also, die von allen Bürgern im Lichte ihrer umfassenden Überzeugungen gleichermaßen anerkannt werden muß. Rawls' Argument für den dichten Veil of Ignorance, der auch die dritte Stufe realisiert, setzt deshalb unzulässigerweise bereits voraus, was mit seiner Hilfe erst hergeleitet werden soll: oberste Grundsätze politischer Gerechtigkeit.

Sobald wir in Übereinstimmung mit Rawls' Konzeption eines *vernünftigen* Pluralismus annehmen (PL: 58 ff.), daß bestimmte Minimalbedingungen der Rationalität und Unparteilichkeit erfüllt sind, müssen wir davon ausgehen, daß umfassende moralische und religiöse Konzeptionen des Guten angemessene und

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

mutmaßlich starke Argumente für die Wahl von Grundsätzen politischer Gerechtigkeit begründen. Wie könnten die Parteien in der Original Position diese Argumente ignorieren, ohne ihren Vertragsschluß für alle irrelevant werden zu lassen, die wie wir vor dem ‚Schleier der Unwissenheit' leben und ihre umfassenden moralischen und religiösen Überzeugungen kennen?[16]

Liegt dem politischen Liberalismus, wie viele seiner Kritiker vermuten, die Vorstellung zugrunde, alle inhaltlichen Konzeptionen eines guten Lebens seien gleichrangig und gleichermaßen ungeeignet, den Inhalt oberster Gerechtigkeitsgrundsätze festzulegen? Benthams ‚pushpin is as good as poetry' steht für diese Auffassung,[17] aber auch bei Rawls lassen sich Stellen finden, die in diese Richtung zu weisen scheinen.[17a] Dies wäre freilich eine schwache Basis, da offen zutage liegende Unterschiede zwischen mehr oder weniger partikularistischen Lebensentwürfen und umfassenden moralischen Lehren bestehen. Letztere treten mit Universalitätsansprüchen auf und artikulieren Konzeptionen des guten Lebens, die potentiell alle Mitglieder einer Gesellschaft umfassen. Die für die Rawls'sche Konstruktion der Original Position entscheidende Frage lautet deshalb: Unter welchen Bedingungen läßt sich der für die deduktive Herleitung der Gerechtigkeitsgrundsätze notwendige Ausschluß umfassender moralischer und religiöser Überzeugungen rechtfertigen, der Ausschluß von Überzeugungen also, denen aus der Sicht der betroffenen Bürger die allergrößte Bedeutung zukommt?

3.4 Eine zunächst Erfolg versprechende Verteidigung des dichten Schleiers der Unwissenheit läßt sich mit Hilfe eines Rational Choice-Arguments entwikkeln, das Allan Buchanan in seinem Aufsatz „Revisability and Rational Choice" von 1975 vorgestellt hat.[18] Sie knüpft an das dritte übergeordnete Interesse

---

16 Hinzu kommt das, was Rawls über die sogenannten „strains of commitment" sagt (TJ: 145, 176 f.). Die Parteien in der Original Position müssen sich ‚in gutem Glauben' auf Gerechtigkeitsgrundsätze einigen und davon ausgehen, daß die von ihnen unter dem Ausschluß bestimmter Informationen gewählten Grundsätze denen keine unzumutbaren Belastungen auferlegen, für die diese Grundsätze im wirklichen Leben verbindlich sein werden. Politische Gerechtigkeitsgrundsätze, die mit den umfassenden moralischen und religiösen Überzeugungen einer Person inkompatibel sind, müssen aber als solche unzumutbare Belastungen angesehen werden. Die Parteien in der Original Position entscheiden sich deshalb für bestimmte Gerechtigkeitsgrundsätze mit dem Vorbehalt, daß diese für niemanden im Lichte seiner umfassenden Überzeugungen inakzeptabel sind. Ob dies der Fall ist oder nicht, kann freilich nicht vom Standpunkt der Original Position aus entschieden werden, sondern nur von denjenigen, die ihre umfassenden Konzeptionen des Guten kennen. Vgl. Rawls' Ausführungen über die drei möglichen Standpunkte, von denen aus ‚Gerechtigkeit als Fairneß' zu beurteilen ist, PL: 27 f.

17 „The Rationale of Reward", (1825).

17a „We shall insure further that particular inclinations and aspirations, and person's conceptions of their good do not affect the principles adopted. The aim is to rule out those principles that it would he rational to propose for acceptance (. . .) only if one knew certain things that are irrelevant form the standpoint of justice." (TJ: 18, dt. Übers.: 36); vgl. PL: 24 zitiert oben auf S. 113.

18 Canadian Journal of Philosophy 5, (1975), S. 395-408.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

moralischer Personen an, ihre Konzeption des Guten unabhängig vom konkreten Inhalt zu schützen und zu fördern. Buchanan geht, wie auch Rawls, von der Überlegung aus, daß die besondere Konzeption des Guten einer Person keine genetisch oder sozial fixierte Disposition ist, in Übereinstimmung mit bestimmten Werten und Normen zu handeln. Neue Erfahrungen und Einsichten können ebenso wie philosophische Reflexion oder religiöse Besinnung dazu führen, daß wir selbst grundlegende Elemente unseres Lebensentwurfs aufgeben und durch andere ersetzen. Die möglichen Revisionen einer Konzeption des Guten können sowohl die für sie konstitutiven Werte und Grundsätze als auch ihre empirisch-deskriptive Komponente betreffen. So werden sich in der Regel im Laufe des Lebens unsere Einschätzungen dessen ändern, was wir unter günstigen Bedingungen für realisierbar halten; und unsere empirischen Vermutungen darüber, wie wir unsere Ziele und Ambitionen am sinnvollsten und wirksamsten umsetzen können, erweisen sich in vielen Fällen als falsch und korrekturbedürftig.

Dabei ist es unwesentlich, ob die von Fall zu Fall nötigen Revisionen in einem objektivierbaren Sinne einen Fortschritt darstellen, etwa zu größerer Rationalität oder Annäherung an ‚die Wahrheit', oder nicht. Wichtig ist allein, daß die resultierenden Veränderungen der eigenen Konzeption des Guten *erstens* aus der Sicht der betroffenen Personen und in Übereinstimmung mit den Geboten praktischer Rationalität notwendig erscheinen und daß sie *zweitens* nur zum Teil der rationalen Kontrolle durch die Betroffenen unterliegen. Jemand, der im Jugendalter feststellt, daß er nicht mehr an den Gott seiner Kinderjahre glaubt und zum Atheisten wird, kann daraus verschiedene gleichermaßen rationale Konsequenzen ziehen, aber er kann die Zehn Gebote nicht länger darum befolgen, *weil* sie die Forderungen Gottes an seine Geschöpfe ausdrücken. Es gehört deshalb zum Begriff einer rationalen moralischen Person, daß jemand sein Handeln nicht nur an gegebenen Grundsätzen und Zielen zu orientieren vermag, sondern auch fähig und bereit ist, normative Vorgaben gegebenenfalls zu revidieren und durch vernünftige Neuorientierungen zu ersetzen. Buchanan faßt diesen Grundgedanken in Form einer Rationalitätsbedingung (R1) zusammen, die fordert: „One ought, *ceteris paribus,* to maintain an attitude of critical revisability toward one's own conception of the good (or life-plan) and of openmindedness toward competing conceptions" (399).

Aufgrund der unvollständigen Kontrolle über die eigenen normativ relevanten Überzeugungen kann niemand mit Gewißheit davon ausgehen, in allen Abschnitten seines Lebens dieselbe Konzeption des Guten zu bejahen. Ein rationaler Akteur wird deshalb bei seinen aktuellen Entscheidungen die Möglichkeit zukünftiger Revisionen seiner Zielsetzungen und Präferenzen in Betracht ziehen. Es wäre schlicht irrational, alle Entscheidungen so zu treffen, als würde sich die eigene Konzeption des Guten niemals ändern, so daß es unnötig wäre, schon in der Gegenwart möglicherweise veränderten zukünftigen Präferenzen Gewicht beizumessen. Dabei setzen wir voraus, daß menschliche Individuen aufgrund

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ihrer numerischen Identität durch die Zeit nicht nur ein rationales Interesse an der Befriedigung ihrer jeweils aktuellen Präferenzen haben, sondern auch an der Befriedigung ihrer zukünftigen Präferenzen interessiert sind, und zwar unabhängig von deren Inhalt. Dies führt Buchanan zu einer zweiten Rationalitätsbedingung (R2): „One ought, *ceteris paribus,* to attempt to provide for the implementation of those new or revised conceptions of the good which one may develop . . ." (402).

R2 entspricht ebenso wie R1 der alltäglichen Lebenserfahrung. Viele Leute legen eine ‚stille Geldreserve' an für den Fall, daß sie später unversehens etwas wünschen oder benötigen werden, ohne daß sie schon wüßten, was es sein wird. Und auch die Theorie rationaler Entscheidungen scheint R2 zu bestätigen; so sprechen Ökonomen seit Keynes von einer rationalen ‚Liquiditätspräferenz' der wirtschaftlich Handelnden bei ungewissen zukünftigen Präferenzen.[19]

Überträgt man Buchanans Prinzip der rationalen Vorsorge (R2) auf die Wahl von Gerechtigkeitsgrundsätzen hinter einem Veil of Ignorance, so ergibt sich folgendes Bild: Ein rationaler Akteur, der zwar seine besondere Konzeption des Guten kennt, aber nichts über seine gegenwärtigen und zukünftigen Lebensumstände weiß — der sich also hinter einem ‚Schleier der Unwissenheit' der ersten und zweiten Stufe befindet —, würde sich so entscheiden, als befände er sich hinter einem dichten ‚Schleier', der auch die dritte Stufe der Informationsbeschränkung realisiert. Da Informationen über persönliche Lebensumstände schon durch den dünnen ‚Schleier' (Stufen Eins und Zwei) ausgeschlossen sind, fehlt ihm jede verläßliche Basis zur Einschätzung der Wahrscheinlichkeit möglicher Veränderungen in der eigenen Konzeption des Guten. Angesichts der Tragweite und Bedeutung möglicher ‚Konversionen' für sein Leben wird er deshalb sicherstellen, daß Gerechtigkeitsgrundsätze gewählt werden, die für ihn akzeptabel erscheinen, *was immer der Inhalt seiner gegenwärtigen oder zukünftigen Lebensvorstellungen sein mag.* Er wird folglich so wählen, als befände er sich hinter einem *dichten* ‚Schleier der Unwissenheit'.

Die von mir hier im Anschluß an Buchanan vorgestellte Rational-Choice-Interpretation des Veil of Ignorance findet ihre Entsprechung in Rawls' Behandlung der Gewissensfreiheit in *A Theory of Justice* (TJ: 205 ff., dt. Übers.: 234 ff.) und in „The Basic Liberties and their Priority" (PL: 310 ff.). Die Parteien in der Original Position entscheiden sich demnach für religiöse Toleranz, weil sie hinter dem Veil of Ignorance nicht sagen können, ob sie zu einer religiösen (moralischen) Mehrheit oder Minderheit gehören und sicherstellen müssen, daß sie unter allen Umständen der von ihnen bejahten Konzeption des Guten folgen können. Rawls bemerkt, daß dies ‚strictly speaking' kein Argument für Gewissensfreiheit sei.

---

[19] Vgl. *J. M. Keynes,* The General Theory of Employment, Interest and Money, Vol. VII der Collected Writings of JMK, (MacMillan, Cambridge University Press, 1936), S. 166 f.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Es gehe vielmehr darum, deutlich zu machen, in welcher Weise „the veil of ignorance combined with the parties' responsibility to protect some unknown but determinate and affirmed religious, philosophical or moral view gives the parties the strongest reasons for securing this liberty" (PL: 311). Ein Argument, das auch solche Bürger überzeugt, die ihre umfassenden religiösen und moralischen Anschauungen kennen, wird aber dann daraus, wenn wir das von Buchanan und Rawls postulierte Interesse moralischer Personen zugrundelegen, ihre möglichen Konzeptionen des Guten unabhängig von deren konkretem Inhalt zu fördern und die zusätzliche wahre Prämisse einfügen, daß niemand vollständige Kontrolle über die von ihm bejahte Konzeption des Guten hat.

3.5 Trotz ihrer intuitiven Plausibilität bleibt die vorgestellte Argumentation unschlüssig. Sie kann nicht zur Rechtfertigung der dritten Stufe des ‚Schleiers der Unwissenheit' dienen. Zwar ist es im allgemeinen richtig, Vorsorge dafür zu treffen, daß nicht nur aktuelle Präferenzen befriedigt werden können, sondern auch noch unbekannte zukünftige. Es ist aber falsch anzunehmen, daß dies für *alle* zukünftigen Präferenzen unabhängig von ihrem konkreten Inhalt gilt. Dies zeigt sich, sobald wir uns vergegenwärtigen, daß Konzeptionen des Guten nicht einfach geordnete Systeme persönlicher Präferenzen sind (obwohl sie in der Regel zu geordneten Präferenzen führen), sondern normative Überzeugungen mit Universalitätsanspruch einschließen. So kann es zu Inkompatibilitäten zwischen verschiedenen Konzeptionen des Guten kommen, die zwischen schlichten Präferenzen nicht möglich sind. Meine starke aktuelle Präferenz für Tee läßt sich ohne Schwierigkeiten damit vereinbaren, daß ich vielleicht im nächsten Jahr nur noch Kaffee trinken möchte. So mag es zweckmäßig erscheinen, schon jetzt für eine Kaffeemaschine zu sparen, anstatt eine weitere Teekanne anzuschaffen. Anders verhält es sich dagegen, wenn ich annehme, in der Zukunft bestimmte meiner *moralischen* Überzeugungen aufzugeben.

Denken wir uns eine strikte, religiös motivierte Abtreibungsgegnerin, die über eine vollkommene Voraussicht über ihr künftiges Leben verfügt und zwei Dinge mit Gewißheit weiß: (1) daß sie ihre religiösen Überzeugungen aus philosophischen Gründen aufgeben wird und (2) daß sie zu einem späteren Zeitpunkt ein starkes Interesse daran haben wird, selber eine legale Abtreibung vornehmen zu können. Die Frage ist, ob dieses Wissen über ihre zukünftigen Überzeugungen und Präferenzen für sie ein rationaler Grund ist, schon jetzt zu wollen, daß sie zu dem späteren Zeitpunkt eine Abtreibung vornehmen lassen kann. Offenbar nicht, solange sie eine strikte Abtreibungsgegnerin ist.

Das bloße Wissen darüber, welche Präferenzen wir in der Zukunft haben werden, impliziert noch keine Wertung der dann von uns präferierten Alternativen. Es konstatiert lediglich das Vorliegen solcher Wertungen in der Zukunft. Sobald wir dagegen nicht nur wissen, welches unsere zukünftigen Präferenzen sind, sondern darüber hinaus ein Interesse an ihrer Befriedigung haben, bewerten wir diese Alternativen, und zwar, wenn wir eine *reine Zeitpräferenz* für die

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Gegenwart ausschließen, genau in der Weise, wie wir es in der Zukunft tun werden. Aber genau diese Bewertung kann eine strikte Abtreibungsgegnerin aus logischen Gründen nicht akzeptieren. Sie ist überzeugt, daß Abtreibungen zu jeder Zeit unangesehen der betroffenen Personen moralisch unzulässig sind, und muß mithin auch die später von ihr selbst gewünschte Abtreibung verurteilen und ablehnen. Dementsprechend wird sie auch kein aktuelles rationales Interesse an der Möglichkeit einer Abtreibung zu dem späteren Zeitpunkt haben. Dies unterscheidet diesen Fall von dem des Tee- oder Kaffeetrinkens; denn jemand kann ein passionierter Teetrinker sein, ohne sich darauf festzulegen, das Kaffeetrinken für alle Perioden seines Lebens zu verurteilen.

Mein Wissen über Präferenzen, die ich unter bestimmten, zur Zeit nicht erfüllten Bedingungen hätte, führt deshalb nur dann zu entsprechenden aktuellen Interessen, wenn diese Präferenzen mit meinen gegenwärtigen normativen Überzeugungen kompatibel sind. Buchanans Übergang von R1 zu R2 ist mithin nicht allgemeingültig: es gibt schlicht kein rational gefordertes Interesse moralischer Personen an der Verfolgung inhaltlich noch unbestimmter Konzeptionen des Guten. Rawls' Begriff der moralischen Person erscheint in dieser Hinsicht korrekturbedürftig. Es ist weder grundsätzlich unangemessen noch irrational, politische Gerechtigkeitsgrundsätze auf der Basis umfassender moralischer und religiöser Überzeugungen zu identifizieren. Bürger, die ihre umfassenden Konzeptionen des Guten kennen, werden mithin — entgegen der Rawls'schen Annahme — nicht schon aufgrund ihres Status als moralische Personen bereit sein, politische Grundsätze allein auf der Grundlage eines übergeordneten Interesses an der Verfolgung jeder möglichen Konzeption des Guten auszuwählen. Wie aber können wir dann rechtfertigen, daß Bürger bei der Erörterung und Auswahl von Gerechtigkeitsgrundsätzen von ihren umfassenden normativen Überzeugungen, insoweit diese nicht allgemein geteilt werden, absehen und sich auf die Informationsbeschränkungen der Original Position einlassen?

4.

4.1 Die überzeugendste Interpretation des dichten Veil of Ignorance dritter Stufe beruht meines Erachtens auf dem Gedanken der öffentlichen Rechtfertigung politischer Grundsätze und dem zugehörigen Postulat eines übergreifenden Konsenses vernünftiger umfassender Lehren. Grundlegend ist der Begriff einer *vernünftigen* Konzeption des Guten. Er führt zu der Idee eines vernünftigen Pluralismus und erlaubt uns eine vergleichsweise einfache Analyse des Problems der öffentlichen Rechtfertigung von Normen in pluralistischen Gesellschaften.

Die grundsätzliche Berechtigung der im Gedanken einer Pluralität *vernünftiger* Lehren gelegenen Idealisierungen ergibt sich daraus, daß wir eine Gerechtigkeitskonzeption suchen, der alle Bürger im Lichte wohlerwogener Überzeugungen aus guten Gründen zustimmen können. Entscheidend für die Rechtfertigung

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

normativer Grundsätze ist offenbar nicht, daß die Bürger sie auf der Basis ihrer faktischen Überzeugungen, wie irrational und unbegründet sie auch sein mögen, *de facto* anerkennen. Worauf es ankommt ist, daß sie nach gebührender Überlegung unter Berücksichtigung aller relevanten Gesichtspunkte anerkannt würden. Die für oder gegen bestimmte Grundsätze vorgebrachten Argumente müssen deshalb einer Reihe von Bedingungen genügen, deren Erfülltsein wir — möglicherweise kontrafaktisch — unterstellen, wenn wir ein Argument als Grund für oder als einen Einwand gegen eine bestimmte Gerechtigkeitsvorstellung gelten lassen.

Dabei gehen wir davon aus, daß Bürger ihre Argumente stets im Rückgriff auf mehr oder weniger umfassende moralische oder religiöse Vorstellungen darüber, was sie für gut und richtig halten, erläutern und begründen. Und wir wissen, daß nicht alle diese Vorstellungen allgemeine Zustimmung finden. Bürger beurteilen politische Gerechtigkeitsgrundsätze auf der Basis divergierender Konzeptionen des Guten, deren je spezifischer Inhalt festlegt, ob ein bestimmter Grundsatz rationalerweise anerkannt werden kann oder nicht. Da nun ein Grund für oder ein Einwand gegen eine bestimmte Gerechtigkeitskonzeption genauso vernünftig oder unvernünftig ist wie die Voraussetzungen, auf denen er beruht, sind die für die öffentliche Rechtfertigung von Normen relevanten Anforderungen vernünftigen Argumentierens zugleich Kriterien der Vernünftigkeit für Konzeptionen des Guten.

Die erste grundlegende Anforderung an vernünftige Konzeptionen und Lehren liegt unter diesem Gesichtspunkt darin, daß sie den allgemein anerkannten Bedingungen theoretischer und praktischer Rationalität genügen. Dazu gehören die Minimalforderungen der logischen Konsistenz und der Übereinstimmung mit allgemein anerkannten Tatsachen. Das schließt auch ein, daß umfassende Konzeptionen des Guten gewissen anthropologischen Grundtatsachen Rechnung tragen und unter den Bedingungen der menschlichen Bedürftigkeit angesichts knapper Ressourcen ein sinnvolles Leben ermöglichen. Konzeptionen, die den Anforderungen theoretischer und praktischer Rationalität uneingeschränkt genügen, können von Bürgern unter Berücksichtigung aller relevanten Tatsachen öffentlich vertreten werden: Auf ihnen beruhende Argumente für oder gegen bestimmte politische Grundsätze können weder als intern unschlüssig noch als empirisch falsch zurückgewiesen werden.

Die allgemeine Rationalitätsforderung erscheint vergleichsweise unproblematisch. Alle Bürger haben normalerweise unabhängig vom konkreten Inhalt ihrer normativen Überzeugungen gute Gründe, sich um die interne Rationalität ihrer Konzeptionen des Guten zu bemühen. Im Kontext der öffentlichen Rechtfertigung von Normen unterstellen wir, daß erkennbar inkonsistente oder empirisch widerlegbare Argumente nach gebührendem Überlegen von ihren Proponenten zurückgezogen würden, weil sie nicht dem entsprechen, was sie unter Berücksichtigung aller relevanten Tatsachen für richtig halten und vertreten wollen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Zu den Grundtatsachen, deren Anerkennung sich aus den allgemeinen Rationalitätsbedingungen ergibt, gehören auch die von Rawls sogenannten ‚Bürden des Urteilens'. Sie *erklären,* warum selbst unter idealen Bedingungen einer uneingeschränkt rationalen moralischen Urteilsbildung nicht ausgeschlossen werden kann, daß Bürger verschiedene moralische oder religiöse Konzeptionen des Guten bejahen. Häufig sind die ein moralisches Problem betreffenden empirischen Befunde uneindeutig, und die offenen Grenzen des korrekten Gebrauchs empirischer und normativer Begriffe lassen in vielen Fällen einen Spielraum für divergierende Interpretationen. Einschlägige normative Gesichtspunkte können mit guten Gründen verschieden gewichtet werden, und miteinander inkompatible moralische Erwägungen erlauben nicht immer eindeutige Problemlösungen. Auch lassen sich nur selten alle moralischen und religiösen Werte gleichermaßen von einer Person oder durch eine Institution verwirklichen. So kommt es zu ‚tradeoffs', auf die es in der Regel mehrere vernünftige Antworten gibt (PL: 54-58). In diesen Bürden des Urteilens treten unüberwindbare Grenzen dafür zutage, durch rationale Argumentation eindeutige Antworten auf alle moralischen und politischen Fragen zu finden. Sie anzuerkennen, bedeutet die Anerkennung einer rational nicht aufhebbaren Pluralität vernünftiger umfassender Lehren und schließt die Bereitschaft ein, die moralischen und religiösen Überzeugungen anderer Bürger, auch wenn sie von den eigenen wohlerwogenen Auffassungen abweichen, als potentiell vernünftig und rational zu betrachten.

Eine zweite, über die allgemeine Rationalitätsbedingung hinausgehende Anforderung an vernünftige Lehren besteht in der Forderung nach Unparteilichkeit. Vernünftige Lehren artikulieren umfassende Konzeptionen des Guten, die wenigstens im Prinzip von allen Bürgern geteilt werden könnten, weil sie auf unparteiischen Werten und Grundsätzen beruhen. Die in einer pluralistischen Gesellschaft bestehenden Divergenzen zwischen moralischen oder religiösen Lehren beruhen nicht auf widerstreitenden partikularen Interessen der Bürger, wie sie sich aus ihren sozialen und ökonomischen Lebensumständen ergeben. Dies schließt nicht aus, daß sich in konkreten historischen Situationen Korrespondenzen zwischen den wirtschaftlichen Interessen gesellschaftlicher Gruppen und den von ihnen vertretenen moralischen oder religiösen Lehren ergeben mögen. Entscheidend ist, daß ein glaubwürdiger Anspruch auf Unparteilichkeit erhoben wird und daß sich die zwischen vernünftigen Lehren bestehenden Divergenzen aus den Schwierigkeiten ergeben, angesichts der Bürden des Urteilens eindeutig zu bestimmen, durch welche materialen Werte und Grundsätze ein unparteiischer moralischer Standpunkt am besten konkretisiert wird.

Für das Problem der öffentlichen Rechtfertigung politischer Gerechtigkeitsgrundsätze ist die Forderung nach Unparteilichkeit deshalb wesentlich, weil die bloße Artikulation partikularer Interessen keine Argumente für oder gegen politische Gerechtigkeitsgrundsätze zu stützen vermag. Mögliche Einwände gegen die in der Original Position gewählten Grundsätze können sich deshalb nur im

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Rahmen vernünftiger Lehren ergeben, die, obwohl sie in einer pluralistischen Gesellschaft *de facto* nicht von allen Bürgern geteilt werden, zumindest einen begründeten Anspruch erheben, im Namen aller zu sprechen.[20]

4.2 Die dritte und neben der Anerkennung der Bürden des Urteilens wichtigste Anforderung an vernünftige Lehren und Konzeptionen des Guten besteht darin, daß die im zweiten Abschnitt vorgestellte kantische Konzeption individueller Autonomie anerkannt wird. Wir können diese Bedingung in Form eines Grundsatzes formulieren, den ich das kantische Autonomieprinzip nennen möchte: *Jeder Bürger hat als freie und gleiche moralische Person einen unabweisbaren Anspruch, nur solchen Normen unterworfen zu werden, denen er auf der Basis der von ihm vertretenen rationalen und unparteiischen Konzeption des Guten*[21] *zustimmen kann.*

Dem kantischen Autonomieprinzip liegen im wesentlichen vier Annahmen zugrunde, die von allen vernünftigen Lehren ebenfalls anerkannt werden: (1) Jeder Bürger ist frei, insofern seine Selbstbestimmung nicht in Weisen eingeschränkt werden darf, die sich nicht auch ihm selbst gegenüber rechtfertigen ließen. (2) Jeder Bürger ist gleich zumindest in dem Sinne, daß politische Institutionen allen gegenüber gerechtfertigt sein müssen. (3) Jeder Bürger ist grundsätzlich fähig und bereit, rationale Überlegungen darüber anzustellen, wie die Institutionen seiner Gesellschaft gerechterweise eingerichtet werden sollten. (4) Jeder Bürger ist grundsätzlich fähig und bereit, faire Beschränkungen seiner ungehinderten Selbstbestimmung anzuerkennen und in seinem Handeln zu beachten.

Personen gegenüber, auf welche die ersten beiden Annahmen (Freiheit und Gleichheit) nicht zutreffen, ist eine Rechtfertigung unnötig; unmöglich ist sie, wenn die für das Verständnis öffentlicher Rechtfertigungen erforderlichen kognitiven Fähigkeiten fehlen, oder wenn jemand wegen der mit allen unparteiischen Regelungen gelegentlich verbundenen persönlichen Nachteile grundsätzlich nicht

---

[20] Welche Bedeutung der Unterscheidung zwischen partikularen moralischen oder religiösen Konzeptionen des Guten auf der einen Seite und partikularen persönlichen oder sozialen Interessen auf der anderen Seite im Selbstverständnis moderner Verfassungsstaaten zukommt, zeigt das Recht auf Kriegsdienstverweigerung, wie es etwa in Deutschland durch das Grundgesetz garantiert wird (Art. 4 (3) und 12a). Niemand kann um privater beruflicher Vorteile willen von der Wehrpflicht befreit werden. Anerkannt werden ausschließlich ‚Gewissensgründe'. Das Grundrecht auf Kriegsdienstverweigerung setzt eine moralisch oder religiös begründete ‚pazifistische' Einstellung des Betroffenen voraus, und dies obwohl die unterstellten normativen Überzeugungen anerkanntermaßen ebensowenig von allen Bürgern geteilt werden wie die grundrechtlich nicht geschützten privaten Interessen.

Zur Forderung der Unparteilichkeit für moralische Lehren im Kontext der öffentlichen Rechtfertigung von Handlungen, Normen und Institutionen vgl. auch *Amy Gutmann* and *Dennis Thompson,* „Moral Conflict and Political Consensus", Ethics 101 (1990), S. 76-86, S. 70 f.; und *Ronald Dworkin,* „Lord Devlin and the Enforcement of Morals", in: Richard A. Wasserstrom, Morality and the Law, Belmont 1971, S. 55-72.

[21] Einer Konzeption also, welche die ersten beiden Anforderungen an vernünftige Lehren erfüllt.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

bereit ist, Gerechtigkeitsgrundsätze ernsthaft und aufrichtig anzuerkennen. Öffentliche Rechtfertigungen richten sich nicht an rationale Egoisten, die Ansprüche anderer nur insoweit berücksichtigen, als dies ihrem eigenen Vorteil dient, sondern an rationale moralische Personen, die einander als freie und gleiche Bürger anerkennen und darum zugestehen, daß politische Institutionen die Interessen und normativen Überzeugungen aller Gesellschaftsmitglieder gleichermaßen achten und berücksichtigen müssen.

4.3 Welche Konsequenzen ergeben sich nun aus dem Gedanken einer Vielfalt vernünftiger umfassender Lehren, die den angeführten Bedingungen genügen, für das Problem der öffentlichen Rechtfertigung politischer Grundsätze?

Im idealtypischen Grenzfall enthalten vernünftige umfassende Lehren weder empirische Fehlannahmen noch Inkonsistenzen. Sie erheben einen begründeten Anspruch auf Unparteilichkeit und erkennen die Bürden des Urteilens ebenso wie das kantische Autonomieprinzip uneingeschränkt an. Alle für öffentliche Rechtfertigungen relevanten Divergenzen zwischen den verschiedenen Konzeptionen des Guten beruhen ausschließlich auf den Bürden des Urteilens und niemals auf einem Mangel an Einsicht oder ungerechtfertigten Parteinahmen. Wir betrachten sie in diesem Sinne als *vernünftige* oder *begründete* Meinungsverschiedenheiten. Eine Situation, in der Bürger vernünftigen umfassenden Lehren anhängen, können wir mit Rawls als vernünftigen Pluralismus bezeichnen (PL: 36 f., 63 ff.).

Das in unserem Zusammenhang wichtigste Merkmal eines vernünftigen Pluralismus besteht in der Asymmetrie zwischen der rationalen Bejahung und der rationalen Verneinung von Konzeptionen des Guten: Jeder Bürger kann unter den beschriebenen idealtypischen Bedingungen für sich reklamieren, seine eigene Konzeption in voller Übereinstimmung mit den Anforderungen der Rationalität und Vernünftigkeit zu vertreten, aber keiner ist in der Lage, die Konzeption irgendeines anderen Bürgers als irrational oder unvernünftig zurückzuweisen. In einer solchen Situation erfordert das kantische Autonomieprinzip, daß keine politische Ordnung auf der Basis von Grundsätzen etabliert wird, die von einer der involvierten vernünftigen Lehren nicht anerkannt wird. Keine umfassende Konzeption des Guten, die Gegenstand begründeter Meinungsverschiedenheiten ist, kann in ihrer Gänze für die öffentliche Rechtfertigung von politischen Grundsätzen und Institutionen in Anspruch genommen werden. Die Rechtfertigung politischer Gerechtigkeitsgrundsätze kann nur mit Hilfe von Prämissen gelingen, über die ein alle vernünftige Lehren übergreifender Konsens besteht.

Daraus ergeben sich von allen vernüftigen Lehren anerkannte Argumentationsbeschränkungen, die ihren angemessenen Ausdruck in der Rawls'schen Konstruktion der Original Position, genauer: im dichten Veil of Ignorance finden. Er ist insofern gerechtfertigt, als die Anerkennung des kantischen Autonomieprinzips unter den Bedingungen eines vernünftigen Pluralismus dazu führt, daß nur solche Argumente den Inhalt der politischen Gerechtigkeit festlegen dürfen, die auch

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

hinter einem dichten ‚Schleier der Unwissenheit', das heißt ohne Inanspruchnahme kontroverser umfassender Konzeptionen des Guten überzeugen.

Nehmen wir dies zusammen mit dem, was oben über den Begriff der moralischen Person gesagt worden ist, so können wir feststellen, daß sich die wesentlichen Bauelemente der Original Position aus dem vertragstheoretischen Gedanken der öffentlichen Rechtfertigung entwickeln lassen. Einwände gegen sie können zurückgewiesen werden, weil sie entweder intern irrational oder mit dem Gedanken der öffentlichen Rechtfertigung unvereinbar sind. Wenn es nun, wie ich durchgängig angenommen habe, zutrifft, daß die Rawls'schen Gerechtigkeitsgrundsätze aus der Original Position deduktiv hergeleitet werden können, dann folgen sie aus Annahmen, gegen die keine vernünftigen Einwände vorgebracht werden können. Wir dürfen insoweit davon ausgehen, daß sie die Unterstützung eines übergreifenden Konsenses finden. Der dichte Veil of Ignorance erscheint unter diesem Gesichtspunkt als eine argumentative Operationalisierung des Gedankens der öffentlichen Rechtfertigung und der mit ihm verbundenen Forderung eines übergreifenden Konsenses für politische Grundsätze.

4.4 Eine aufschlußreiche Komplikation dieser Interpretation des Veil of Ignorance ergibt sich nun daraus, daß das kantische Autonomieprinzip vernünftige, aber kontroverse Konzeptionen des Guten zwar als Grundlage zur Rechtfertigung politischer Grundsätze ausschließt, sie als Basis für öffentlich anzuerkennende Einwände aber ausdrücklich zuläßt. Jede kontroverse, aber für die Rechtfertigung einer Norm nötige Prämisse begründet einen vernünftigen Einwand gegen die Norm und hebt nach dem kantischen Autonomieprinzip deren Allgemeinverbindlichkeit zwingend auf. Wie leicht zu sehen ist, ist eine Prämisse A zur Rechtfertigung einer Norm N! genau dann unverzichtbar, wenn ‚N! → A' gilt, so daß wegen ‚nicht-A → nicht-N' jeder vernünftige Zweifel an A einen Zweifel an N! begründet. Im Kontext der öffentlichen Rechtfertigung von Normen und Institutionen können demnach kontroverse normative oder faktische Annahmen wenn nicht als *Argumente für,* so doch als *Einwände gegen* bestimmte Grundsätze in Anspruch genommen werden.

Die öffentliche Rechtfertigung von politischen Grundsätzen ist demnach noch nicht gelungen, wenn gezeigt ist, daß diese aus Annahmen folgen, in bezug auf die alle vernünftigen Lehren übereinstimmen. Gefordert ist darüber hinaus, daß sich auf der Grundlage keiner der involvierten vernünftigen Lehren ausschlaggebende Einwände ergeben. Wegen des kantischen Autonomieprinzips müßten diese Einwände auch dann öffentlich anerkannt werden, wenn sie auf kontroversen Annahmen beruhen, vorausgesetzt nur, sie erfüllen die angegebenen Kriterien der Rationalität und Vernünftigkeit. Wie läßt sich dies aber mit dem dichten Veil of Ignorance vereinbaren, der alle kontroversen Überzeugungen aus der Sphäre öffentlicher Argumente ausschließt?

Wegen der unvermeidlichen Fallibilität aller praktischen Argumente kann grundsätzlich nicht ausgeschlossen werden, daß gegen jede denkbare politische

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Gerechtigkeitskonzeption vernünftige Einwände vorgebracht werden. Vergegenwärtigen wir uns jedoch die argumentative Situation, in der sich jemand befindet, der auf der Basis einer vernünftigen Konzeption des Guten versucht, einen begründeten Einwand gegen die in der Original Position angenommenen Gerechtigkeitsgrundsätze zu erheben. Unproblematisch sind Einwände, die von allen Bürgern unangesehen ihrer besonderen Konzeptionen des Guten als vernünftig und zutreffend anerkannt werden. Sie stellen die Möglichkeit einer öffentlichen Rechtfertigung politischer Grundsätze nicht in Frage; denn sie können, eben weil sie keine besondere Konzeption des Guten voraussetzen, auch hinter einem dichten Veil of Ignorance von den Parteien in der Original Position berücksichtigt werden.

Der interessantere Fall liegt vor, wenn Bürger auf der Basis einer vernünftigen, aber umstrittenen Konzeption des Guten Einwände erheben — Einwände also, die von allen Bürgen unangesehen ihrer besonderen Konzeption des Guten als vernünftig anerkannt werden, die aber nicht von allen geteilt werden.[22] Man beachte zunächst, daß wir es in einem solchen Fall nicht nur mit einem Konflikt zwischen einer besonderen Konzeption des Guten und politischen Gerechtigkeitsgrundsätzen zu tun haben, sondern zugleich mit einem internen Konflikt innerhalb der betreffenden Konzeption des Guten. Da wir annehmen, daß die materialen Prämissen der Herleitung der beiden Grundsätze von allen vernünftigen Lehren akzeptiert werden, sprechen aus der Perspektive jeder dieser Lehren gewichtige Gründe für ihre Anerkennung, und zwar unabhängig davon, welche Einwände darüber hinaus *gegen* die Grundsätze zu berücksichtigen sein mögen. Entscheidend ist deshalb das relative Gewicht der zu berücksichtigenden Einwände. Auf der Grundlage einer vernünftigen Konzeption des Guten ergibt sich nur dann ein triftiger Einwand gegen bestimmte Gerechtigkeitsgrundsätze, wenn die Gründe, von denen wir wissen, daß sie im Rahmen jeder solchen Konzeption für die Grundsätze sprechen, durch gewichtige entgegenstehende Gründe überwogen werden.

Meines Erachtens sprechen nun mindestens drei Überlegungen dafür, daß sich im Rahmen vernünftiger Konzeptionen des Guten keine *ausschlaggebenden* Einwände gegen die in der Original Position gewählten Grundsätze ergeben. *Erstens* wird der der Original Position zugrundeliegende Personbegriff von allen vernünftigen Lehren anerkannt, und niemand bestreitet vernünftigerweise, daß das Interesse aller Bürger an der adäquaten Ausbildung ihrer beiden moralischen Vermögen für die Wahl von Gerechtigkeitsgrundsätzen überragende Bedeutung hat.

---

[22] Solche Einwände müssen drei Bedingungen erfüllen: (1) Unter der Voraussetzung, daß man ihre Prämissen bejaht, lassen sie eine Revision der Grundsätze, gegen die sie gerichtet sind, notwendig erscheinen. (2) Die in Anspruch genommenen Prämissen können in voller Übereinstimmung mit den Anforderungen der Rationalität und Vernünftigkeit sowohl bejaht als auch verneint werden. (3) Nicht alle Bürger bejahen die fraglichen Prämissen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Zweitens wird der große Wert friedlicher sozialer Kooperation auf der Basis gegenseitiger Achtung, wie sie durch die Rawls'schen Grundsätze gewährleistet wird, nicht leicht überwogen. *Drittens* schließt das von allen vernünftigen Lehren anerkannte Autonomieprinzip kontroverse Annahmen als positive Grundlage einer politischen Gerechtigkeitskonzeption aus. Und dies legt nahe, daß Vertreter vernünftiger Lehren Einwänden, die auf umstrittenen Annahmen beruhen, ein geringeres Gewicht beimessen als den von allen anerkannten normativen Konzeptionen, die der Original Position zugrundeliegen. So wird etwa der Vorrang der Gewissensfreiheit selbst angesichts eines rationalen Interesses an religiöser Homogenität von keiner vernünftigen Lehre in Frage gestellt, weil sich eine entsprechende Einschränkung der bürgerlichen und politischen Freiheiten nicht gegenüber ‚Andersgläubigen' rechtfertigen ließe.

Ungeachtet dieser Überlegungen, so zutreffend sie sein mögen, kann die Möglichkeit vernünftiger, aber kontroverser Einwände gegen oberste politische Gerechtigkeitsgrundsätze nicht grundsätzlich ausgeschlossen werden, und es ist müßig, darüber zu spekulieren, wie wahrscheinlich sie sind. Werden entsprechende Einwände vorgebracht, so ist zu überprüfen, welche Konsequenzen sich im einzelnen mit Blick auf das von allen geteilte Ziel ergeben, eine öffentlich gerechtfertigte Gerechtigkeitskonzeption zu finden. Festzuhalten ist jedenfalls, daß die Offenheit gegenüber vernünftigen Einwänden keine Besonderheit der Rawls'schen Gerechtigkeitsgrundsätze und des für sie reklamierten übergreifenden Konsenses darstellt. Es trifft auf alle praktischen Argumente zu, daß wir ihre Konklusionen stets unter der möglicherweise kontrafaktischen Annahme akzeptieren, daß alle relevanten Gesichtspunkte berücksichtigt sind, und mit dem Vorbehalt, daß keine triftigen Einwände vorgebracht werden. Dies gilt für den von Rawls postulierten übergreifenden Konsens ebenso wie für jede andere Form der rationalen Zustimmung zu praktischen Geboten.

4.5 Hinsichtlich der Bedeutung des dichten Veil of Ignorance für die öffentliche Rechtfertigung politischer Gerechtigkeitsgrundsätze ergibt sich damit folgendes Ergebnis: Als eine argumentative Operationalisierung des Gedankens der öffentlichen Rechtfertigung stellt er sicher, daß die Parteien in der Original Position angesichts eines vernünftigen moralischen und religiösen Pluralismus nur solche Voraussetzungen für die Identifikation oberster politischer Gerechtigkeitsgrundsätze in Anspruch nehmen, die sich auf einen übergreifenden Konsens vernünftiger umfassender Lehren stützen können. Zu Recht schließt der Veil of Ignorance außerdem alle Einwände und Überlegungen aus, die den angegebenen Kriterien praktischer Rationalität und Vernünftigkeit nicht genügen. Mit Blick auf die Möglichkeit vernünftiger, aber kontroverser Einwände, die von ihm ebenfalls ausgeschlossen werden, obwohl sie nach dem kantischen Autonomieprinzip öffentlich anerkannt werden müssen, können wir den Veil of Ignorance als Ausdruck des mit allen praktischen Argumenten verbundenen und stets falliblen Anspruchs verstehen, alle relevanten Gesichtspunkte seien in angemessener Weise berück-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

sichtigt worden. Ob dies tatsächlich der Fall ist oder nicht, muß sich angesichts jedes kontroversen, aber begründeten Einwandes von neuem erweisen.

Ein weiteres für die vertragstheoretische Konzeption der Gerechtigkeit als Fairneß wichtiges Ergebnis meiner Überlegungen liegt darin, daß wir uns nicht darauf beschränken müssen, zu *hoffen,* für oberste Grundsätze politischer Gerechtigkeit einen übergreifenden Konsens zu finden. Ein solcher Konsens kann vielmehr innerhalb der beschriebenen Grenzen durch rationale Argumentation herbeigeführt werden. Dies geschieht mit Blick auf die Rawls'sche Konzeption, indem gezeigt wird, inwiefern sich die Konstruktionselemente der Original Position aus dem Gedanken der öffentlichen Rechtfertigung politischer Grundsätze unter den Bedingungen eines vernünftigen Pluralismus ergeben.

An dieser Stelle mag sich bei manchem der Verdacht regen, die von mir beschriebene Argumentationskette vom vernünftigen Pluralismus hin zu inhaltlich bestimmten politischen Grundsätzen sei zirkulär, da alles schon in den Begriff *vernünftiger* umfassender Lehren und Konzeptionen des Guten hineingelegt worden sei. Dazu drei abschließende Bemerkungen:

Es sollte niemanden überraschen, daß bei deduktiv konzipierten Argumenten alles schon in den Prämissen enthalten ist. Die Kritik kann sich also allenfalls auf eine ungenügende Begründung der in Anspruch genommenen Voraussetzungen beziehen. Namentlich müßte gezeigt werden, daß die für alles andere grundlegende Konzeption eines vernünftigen Pluralismus bereits die gesamte Rawls'sche Gerechtigkeitskonzeption inklusive der beiden Grundsätze voraussetzt. Dies ist offensichtlich nicht der Fall. Ich habe mich bemüht zu zeigen, inwiefern sich die einschlägigen Anforderungen an vernünftige Lehren und Konzeptionen des Guten aus dem allgemeineren Gedanken der öffentlichen Rechtfertigung ergeben, wenn wir fragen, welchen Anforderungen rationale Argumente für oder gegen politische Gerechtigkeitsgrundsätze genügen müssen.

Es wird auch nicht behauptet, daß jede vernünftige Konzeption des Guten für sich genommen bereits die Rawls'schen Gerechtigkeitsgrundsätze impliziert. Dies gilt vielmehr nur unter der Bedingung des Bestehens einer Pluralität solcher Konzeptionen. In einer religiös homogenen Gesellschaft ist es, jedenfalls im Prinzip, denkbar, daß über religiöse Grundsätze politischer Gerechtigkeit keine vernünftigen Meinungsverschiedenheiten bestehen. Ob dieser Fall jemals eingetreten ist oder nicht, ist eine andere Frage. Ich neige dazu, es zu bezweifeln. Dies führt uns auf die schwierige Frage, welche konkreten Bedingungen erfüllt sein müssen, um mit Blick auf historische Beispiele von einem vernünftigen Pluralismus zu sprechen. Eine Schwalbe macht noch keinen Sommer, aber schon ein Häretiker hat, wenn man dem Autonomiegedanken folgt, einen unabweisbaren Anspruch, nur Regeln unterworfen zu werden, gegen die auch er keine vernünftigen Einwände erheben kann.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Schließlich möchte ich hervorheben, daß es nicht verwunderlich ist, wenn wir unter den Bedingungen eines vernünftigen Pluralismus eine Konvergenz aller vernünftigen Lehren in bezug auf bestimmte Gerechtigkeitsvorstellungen feststellen können. Alle Lehren befinden sich ja, was die Rechtfertigung politischer Grundsätze angeht, in derselben normativen Situation. Das Faktum des Pluralismus ist nicht nur ein Problem für die politische Philosophie, sondern jede moralische und religiöse Lehre muß es als normative Grundtatsache berücksichtigen, und erst darin, wie dies geschieht, erweist sich ihre Vernünftigkeit.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Rationalität, Moral und Emotion

Von *Birger P. Priddat,* Witten

Ökonomische Rationalität und Emotionalität schließen sich scheinbar aus. Die Moral steht dazwischen. Meist wird sie — als ‚moralisches Gefühl' — den Emotionen zugeschlagen. Folglich wird die Moral dann genauso von der Rationalität distanziert wie die Emotion. Oder aber es wird nur diejenige Moral für verhandlungsfähig erklärt, die selber auf rationaler Grundlage steht — „Passions within Reason".[1]

Doch ist die Lage der Dinge nicht so einfach. Es läßt sich die These vertreten, daß die Ökonomie dort, wo sie auf rationales Verhalten gründet, immer bereits schon Emotionen verarbeitet. Eine andere, zweite These lautet: Die rational choice-Theorie baut implizit auf einer moralischen Grundlage auf. Beide Thesen harmonieren nicht. Es wird sich zeigen lassen, daß die moralische Basis der rational choice eine ‚vernunfttheoretische' Konstruktion ist, die dazu dient, die Emotionen *normativ* aus dem Entscheidungsverhalten auszuschließen, während die Emotionen — in den Wünschen etc. — auf der anderen Seite *positiv* die Basis der rationalen Entscheidungsfigur darstellen.

Zwischen beiden Komponenten der Theorie besteht eine Spannung, die nicht rational aufgelöst werden kann. Wenn sie dann aber auch nicht emotional aufgelöst werden soll, bleibt nur die Moral, deren systematischer Ort in der Ökonomie neu bestimmt werden kann. Hierzu will der folgende Beitrag einige Anregungen geben.

1. Die Emotionen verarbeitende rational choice

Die Wünsche, Erwartungen, Antriebe etc., die den rational actor überhaupt erst zum Handeln anleiten, sind mehr oder minder emotionale Tensoren; sie erzeugen die Spannung ihrer Realisation. Dabei wird schlicht davon ausgegangen, daß Wünsche etc. meistens emotionale Zustände der Akteure sind, jedenfalls undeutlich nur von Leidenschaften, Affekten etc. zu trennen.[2]

---

[1] So der Titel von *Robert Franks* einschlägigem Buch (N.Y. / London 1988). Im deutschen Titel „Strategie der Emotionen" (München 1992) wird diese Implikation nicht offensichtlich.

[2] Dieser Umstand birgt für theoriegeschichtlich geschulte Ökonomen keine Überraschung, ist doch der Bentham'sche Utilitarismus eine Konzeption der Maximierung von

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Erst in der Konfrontation mit den (knappen) Handlungsalternativen entsteht an den Wünschen etc. ein Wahlproblem. Was man sich wünscht, möchte man auch haben. Zwei Restriktionen behindern allerdings eine ‚paradiesische' Lösung: Zum einem gibt es sehr viele Wünsche, die miteinander konkurrieren, d.h. nicht alle gleichzeitig verwirklicht werden können. Zum anderen kann ein Wunsch nur dann erfüllt werden, wenn die Mittel zu seiner Realisation vorhanden sind. Beide Möglichkeiten indizieren Knappheiten.

Unter Knappheitsbedingungen muß man sich entscheiden, was allerdings nicht automatisch zu einem rationalen Wahlvorgang führen muß. Die Knappheit von Handlungsalternativen kann auch durch einfache Entscheidungen (Dezisionen) bewältigt werden.[3] ‚Knappheit' ist nur eine andere Redeweise für den Umstand, daß mehr Wünsche vorgebracht werden, als aktuell realisiert werden können. Rationale Wahlhandlungen dagegen zeichnen sich *zusätzlich* durch eine besondere Entscheidungsform aus: Daß die beste der gegebenen Alternativen gewählt werden soll.[4]

Die Zusätzlichkeit der Rationalitätsbedingung von Entscheidungen ist nicht von vornherein ersichtlich. Sie ist keine — wie man meinen könnte — anthropologische Aussage über ‚den Menschen', sondern nurmehr eine Aussage über den Umstand, daß nicht alle Alternativen, die ein Individuum für zweckmäßig hält, erlangbar sind, weil ökonomisches Handeln immer ein Handeln unter Restriktionen ist: Die Wünsche, Verlangen, Begehren etc. mögen zwar groß sein, aber entweder gibt es eine persönliche Budgetrestriktion, oder aber das Angebot ist begrenzt. „Ökonomie", können wir eine sehr allgemeine Definition anführen, „ist die Untersuchung menschlichen Handelns, insofern es ein Wählen zwischen Alternativen beinhaltet, das Restriktionen unterliegt und von Zwecken geleitet ist" (*Vilks* 1991, S. 1).

Die Zwecke geben an, daß nur das gewählt wird, was ein Individuum für sich als zweckmäßig ansieht; die Restriktionen erklären, warum zwischen Alternativen, die ein Individuum alle gleichermaßen als zweckmäßig ansehen kann, den-

---

Glück, Freude, Lust (‚pleasure') etc. So lesen wir es noch bei Jevons und bei Gossen. Erst seit der ‚Glücksutilitarismus' im 20. Jahrhundert durch den ‚Präferenzutilitarismus' abgelöst wurde (vgl. dazu: *Birnbacher* 1990, S. 72 ff.), der allein die Befriedigung (beliebiger) individueller Interessen zum Gegenstand hat, ist die emotive Komponente verschwunden (vgl. z.B. bei *Harsanyi* 1982). Redet man von ‚Präferenzrealisation', haben wir eine affekt-neutrale Semantik hergestellt, reden wir aber von ‚Bedürfnisbefriedigung', haben wir im Terminus der ‚Befriedigung' immerhin noch einen ‚Gefühls-Rest'.

3 Zur Theorie der Dezisionen bzw. dezisionalen Entscheidungen vgl. das Heft „Ethik und Dezisionismus" (hg. von *M. Kaufmann*) der Zeitschrift „Ethik und Sozialwissenschaften", H. 1, 1992 (darin besonders der Artikel von *Kersting* 1992).

4 Dieser Satz charakterisiert die Struktur aller rational choice-Theorien: Es geht immer um eine eindeutige Entscheidung, die allen anderen möglichen Entscheidungsalternativen vorgezogen wird — gleichgültig, ob in der Form einer Maximierung, ihrer Umkehrung in der Minimierung (auch der Minimierung des ‚Bedauerns') oder in der Optimierung (vgl. dazu *Vilks* 1991, Kap. 1).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

noch zu wählen ist, wenn wir unterstellen wollen, daß es nicht alle Zwecke gleichzeitig zu realisieren vermag. In diesem Sinne ist die ‚rationale Wahl' ein logisches Desiderat zwischen Zwecken und Restriktionen. Man kann nur das wählen, was die (Budget- oder Angebots-)Restriktionen zu wählen erlauben. Das ist allemal weniger als an zweckmäßigen Alternativen zur Wahl steht. Folglich ist es ‚rational', unter den Restriktionen diejenige Alternative zu wählen, die den Zweck am besten erreichen läßt.

Daran zeigt sich, daß die rational choice einen *normativen* Anspruch an das Entscheidungsverhalten hat. Selbst wenn wir es für ein empirisches Faktum halten, daß sich ein Individuum in einer Situation knapper Entscheidungsalternativen für *eine* davon entscheiden muß (wenn es diesbezügliche Wünsche etc. hat), ist es immer eine normative (zusätzliche) Bestimmung, daß es diese Entscheidung rational ausführen soll, indem es ‚die beste' Alternative wählt. Das Rationalitätsprinzip der modernen Ökonomie ist kein notwendiger Bestandteil des Entscheidungsverhaltens (vgl. *Fusfeld* 1992, S. 62). Denn es reichte für die Allokation der knappen Ressourcen, wenn die Individuen sich für irgend*eine* der Alternativen entschieden.[5]

Das Rationalitätspostulat ist somit keine notwendige Bedingung, sondern eine zusätzliche Anforderung an das Allokationsverfahren: Daß die Akteure nicht irgendwelche Entscheidungen tätigen, sondern diejenige (singuläre) Entscheidung, die sie — in Kenntnis aller anderen Entscheidungsmöglichkeiten, die sie auch hätten wählen können — *nicht mehr revidieren wollen.* Die Akteure sollen sich nicht nur einen der vielen Wünsche etc. erfüllen, die sie haben, sondern sicher gehen können, daß sie sich aktuell den Wunsch erfüllen, den sie am stärksten verspüren. Das aber ist eine höchst problematische Anforderung, denn die Wünsche der Individuen stehen — ohne weitere Qualifikation — in emotionaler Konkurrenz.[6]

---

[5] Man kann sogar sagen, daß die Allokation knapper Ressourcen auf der Basis dieser dezisionalen Entscheidungen ‚optimal' ist. Der Allokationsprozeß kann keine anderen Entscheidungen verarbeiten als die, die getätigt wurden. Im Rahmen dieses ‚Inputs' wird so alloziiert, als ob den Akteuren keine anderen Entscheidungen möglich gewesen wären: Das Ergebnis ist — unter diesen Bedingungen — ‚optimal', da — empirisch — kein anderes entstehen konnte. Alle anderen Aussagen bedürfen einer normativen Qualifikation, z. B. dann, wenn man behauptet, daß das Gesamtergebnis hätte ‚besser' sein können. Man muß dann zeigen können, daß die Akteure sich ‚besser' hätten entscheiden können, als sie es faktisch getan haben. Wenn die Akteure aber behaupten, daß sie sich so entscheiden wollten, kann der normative Anspruch nichts ausrichten: Erst wenn sie selber zweifeln, ob sie es sich nicht ‚besser' hätten überlegen können, gilt die normative Kritik des Allokationsergebnisses.

[6] Zudem ist die Entscheidung, eine Alternative vor allen anderen als die ‚beste' auszuzeichnen, nur normativ als Abschluß des Entscheidungsvorganges festzuschreiben; empirisch gibt es Gründe, gerade dann, wenn man meint, sich auf eine ausgezeichnete Alternative festgelegt zu haben, in Zweifel darüber zu kommen, ob sie die richtige sei. Die Entscheidung selbst produziert neue Unterscheidungen und damit neue Unsicherheit (vgl. auch *Hauschildt* 1988, 77 über „Zielvariation" in Entscheidungsprozessen).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Das Rationalitätspostulat fordert, daß eine Präferenzordnung formuliert werden kann, die nichts anders darstellt als eine *‚Ordnung der Gefühle'*, die die Akteure überhaupt erst in der Lage versetzt, eine Alternative vor allen anderen zu bevorzugen (um nicht durch die spontane Revision der Wünsche selbst irritiert zu werden). Wir haben es mit einem Sublimationsvorgang zu tun, einer Art Bändigung der Leidenschaften, die sie in geordnete Begehrensstrukturen überführt. Diese Ordnung der emotionalen Wünsche etc. kann als Rationalisierungsvorgang der Leidenschaften angesehen werden, dessen formale Form: die rational choice, allerdings heute vollständig in eine kognitive Operation transformiert worden ist.[7] Damit werden ‚voluntative oder leidenschaftliche Entscheidungen' (Dezisionen) ausselegiert, indem sie a priori als ‚irrational' gelten.

Doch ist dies eine unzulässige Einengung. In der Theorie der rationalen Wahlhandlungen können wir zwischen kognitiven, intellektuellen und emotiven Vorgängen nicht deutlich unterscheiden. Wenn zwischen zwei Alternativen zu wählen ist, reicht es für die ‚rationale' Bezeichnung des Vorganges der Wahl vollständig aus, daß eine Alternative vor der anderen präferiert wird. Wie die Entscheidung vonstatten ging — ob sie intellektuell, kognitiv oder affektiv getroffen wurde —, ist theoretisch unerheblich. Und doch wird implizit unterstellt, daß der Wahlakt kognitiv oder intellektuell vonstatten geht, weil nur dann gesichert werden könne, daß gleiche Alternativen gleiche Präferenzen erzeugen. Affektive Wahlen würden dies kontingent oder offenlassen müssen.

Diese Konzeption läßt sich nicht aufrechterhalten. Es gibt keine gleichen Alternativen, und zwar nicht nur, weil es nicht dieselben Alternativen sein können, sondern weil Alternativen keine ‚objektiven' Gegebenheiten sind, sondern subjektive Interpretationen.[8] Folglich nützt es nichts, kognitive oder intellektuelle Prozesse vor den affektiven zu bevorzugen, da die kognitiven oder intellektuellen Prozesse nicht durch Selbstauszeichnung von den affektiven oder emotionalen zu trennen sind. Die Alternativeninterpretation wird durch eine kognitive oder intellektuelle Wahrnehmungsweise nicht stabilisiert, da wir *Entscheidungsstimmungen* nicht ausschließen können (*Priddat* 1994, Kap. 8a).

---

[7] Das letztere wird unmittelbar einsichtig, wenn wir die bayesianische expected utility-Konzeption ansehen, in der erwartetete, d.h. erst zukünftig zu realisierende Nutzen durch subjektive Eintrittswahrscheinlichkeitsannahmen repräsentiert werden.

Es gibt logisch keinen Grund, die Wahrscheinlichkeitsmaße nicht emotional oder intuitiv zu bestimmen; aber die bayesianische rational choice-Konzeption versucht mit Hilfsannahmen über Informationsprozesse die intuitive Charakteristik der Wahrscheinlichkeitsbildung in eine kognitive zu verwandeln. Das ist legitim und in manchen Fällen sinnvoll, aber man muß sich bewußt bleiben, daß wir es mit einer Konstruktion zu tun haben, die nicht das tatsächliche Verhalten der Individuen abbildet, sondern ihnen eine bestimmte kognitive Operation zuschreibt, um auf dem Wege der Rekonstruktion eine rationale Entscheidungsgenese sicherzustellen. Die Konstruktion aber ist dafür verantwortlich, daß die empirischen Ergebnisse der Untersuchung von Entscheidungsverhalten durchaus andere Charakteristiken aufweisen (vgl. *Kahnemann / Slovic / Tversky* 1982).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Wir können zwar normativ behaupten, daß es besser sei, nicht-emotionale Interpretationen zu bevorzugen, aber was nützt dieser Vorschlag, wenn bestimmte kognitive oder intellektuelle Schemata selber affektiv besetzt sind (z.B. dann, wenn wir rationale vor ‚bloß emotionalen' Entscheidungen bevorzugen). Da wir Stimmungen nicht leugnen können, können wir auch nicht ausschließen, daß sie unsere kognitiven oder intellektuellen Repertoires beeinflussen.[9] Somit können wir auch für die rationale Wahl emotionale Einflüsse (oder Determinationen) nicht von der Hand weisen. Unter Anerkennung der Tatsache, daß alle unsere Handlungen emotional gestimmt sind oder von Stimmungen begleitet werden, können wir die Abwehr der Geltung des Emotionalen selber nur als eine Emotion deuten, die keiner Begründung gegen das Faktum ihrer Geltung fähig ist. Uns bleibt nichts anderes übrig, als die Präferenzen als facta anzunehmen, gleichgültig, welcher Stimmung, Überlegung oder Prinzipienhaftigkeit sie entspringen.

Wenn wir aber akzeptieren können, daß bestimmte kognitive Schemata und bestimmte intellektuelle Konzeptionen emotionalen Besetzungen unterliegen können, dann können wir nicht ausschließen, daß die Vorstellung, rational zu handeln, einem Gefühl für Ordnungen ebenso entsprechen kann wie einem intellektuellen oder kognitiven Ordnungsvorgang von Gefühlen. Die Vermutung geht dahin, daß die Präferenzordnung als Ordnung der Gefühle glaubhaft ist[10], weil sie einem Gefühl von Ordnung entspricht, und zwar auch dann, wenn wir gar keinen bewußten Ordnungsprozeß eingeleitet und durchgangen haben. Wir wissen nicht, wie Präferenzordnungen zustande kommen: Gleichwertig können sie durch erwägende Urteile über den intensivsten Wunsch zustande gekommen sein, wie aber auch durch die Dominanz eines Gefühls vor anderen. Die formelle Anforderung nach Transitivität der Präferenzordnungen kann selbst durch ein Gefühl repräsentiert werden, diese und keine andere Ordnung gelten lassen zu wollen. Doch braucht dies wiederum nur für den idealen Moment der Befragung zu gelten; im nächsten Moment gilt vielleicht schon etwas anderes — und zwar für die gleiche Entscheidungssituation.[11] Die Stabilität von Präferenzordnungen ist empirisch nicht haltbar.

Alle Ordnungen, die für die rational choice vorausgesetzt werden (entweder die Vorordnung der Gefühle, Wünsche etc. in einer stabilen Präferenzordnung

---

[8] Vgl. dazu *Lachmann* 1991. Das gilt selbst für die Fälle definiter Güterangebote. Wir wählen nicht zwischen allen Angeboten, z.B. Automobilen, sondern selegieren bereits immer schon bestimmte Typen, Klassen oder Sorten aus, denen wir eine bestimmte ‚Bedeutung' geben. Letztlich wählen wir nur zwischen Alternativen, die wir für bedeutsam halten. Da wir aber nicht alle möglichen Alternativen für bedeutsam halten, wählen wir, bevor wir rational auswählen, die Menge der relevanten Alternativen irgendwie vorher aus (vgl. dazu: *Schütz* 1971; *Esser* 1991).

[9] Vgl. dazu die erhellende Theorie der ‚Affektlogik' von: *Ciompi* 1993a und 1993b.

[10] Vgl. auch *L. A. Berger:* „Utility functions should be called emotional functions" (*Berger* 1991, 271).

[11] Vgl. die untersuchten Fälle in: *Sugden* 1991; auch: *Elster* 1990; *Davidson* 1990b; *Fusfeld* 1992.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

oder aber die resultierende Präferenzordnung, für die es lediglich wichtig ist, die Spitze zu bestimmen, während der Rest ungefähr geordnet bleiben kann, d.h. gegebenenfalls rest-transitiv oder rest-indifferent), beruhen auf der Zuschreibung kohärenter Einstellungen: „Ich glaube . . . nicht, daß wir klar anzugeben vermögen, was uns davon überzeugen würde, daß jemand zu einem bestimmten Zeitpunkt (ohne sich anders zu besinnen) *a* höher als *b, b* höher als *c* und *c* höher als *a* einstuft. Der Grund für unsere Schwierigkeit ist, daß wir Präferenzuzuschreibung nicht verständlich machen können, ohne sie vor einem Hintergrund kohärenter Einstellungen zu sehen" (*Davidson* 1990 (I), S. 322).

Das Faktum, daß die gleichen Entscheidungen zu verschiedenen Zeitpunkten eine intransitive Präferenzordnung anzeigen können (vgl. Elster 1983, Sugden 1991, Anand 1993), widerspricht unserem Kohärenzempfinden, weshalb wir der normativen Theorie der Rationalität (und ihren Axiomen) zugänglich sind, ohne daß diese Zugänglichkeit etwas über das tatsächliche Entscheidungsverhalten aussagt, sondern nur etwas über erwünschte logische Evidenz. Die ‚Logik der Rationalität' entspricht unseren intellektuellen Erwartungen, was mit dem faktischen, emotional getönten Entscheidungsverhalten nichts zu tun zu haben braucht. Im nächsten Abschnitt wird erörtert werden, wie dieses Kohärenzempfinden in der normativen rational choice-Theorie eine moralische Konfundation erfährt.

Damit haben wir das triviale, aber wenig beachtete Resultat, daß die rational choice-Theorie eine Theorie der Ordnung von Emotionen darstellt. Die rational choice steht auf ihrer input-Seite jedwelchen Emotionen offen. Daß dieses Resultat unbeachtet bleibt, liegt schlicht daran, daß vornehmlich der ordnende Charakter der Rationalität erörtert wird, gleichgültig, was sie ordnet.

2. Die der rational choice implizite Moral

Wie die Präferenzordnung gebildet wird, ist für die ökonomische Theorie irrelevant. Das ist der logische Grund, weshalb sie völlig ohne Psychologie — oder mit beliebig vielen Psychologien — auskommen kann. Es genügt ihr vollständig, die in den individuellen Wahlakten effektiv wirksam gewordene Präferenzordnung als ein fait accompli zu nehmen (*Crouch* 1979, S. 4; *Nicolaides* 1990) und *das Individuum für seine Entscheidung zur Verantwortung zu ziehen.*

Dies wird sofort offensichtlich, wenn wir den Entscheidungszusammenhang anders beschreiben. In der gleichgewichtstheoretischen Formulierung des Allokationsoptimum wird normativ unterstellt, daß alle beteiligten Akteure für ihre individuellen Entscheidungen *verantwortlich* sind. Das heißt nichts anderes, als daß *das Gleichgewicht ein Zustand ist, in dem kein Akteur mehr behaupten kann, etwas anderes zu wollen als das, was er entschieden hat.* Die normative Diktion wird offensichtlich: *Es geht nicht darum, wie sich die Akteure tatsächlich verhalten haben, sondern um den Umstand, daß sie, gleichgültig wie sie sich verhalten*

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

haben, ihre eigene Entscheidung vor sich selbst als endgültig und irreversibel zu verantworten haben.

Diese Rechtfertigungsfigur enthält — unerwarteterweise — ein moralisches Argument. Indem normativ unterstellt werden kann, daß alle Akteure rational gehandelt haben — gleichgültig, wie sie sich empirisch tatsächlich verhalten[12] —, kann der Markt als ein optimales Allokationssystem erscheinen, dessen Ergebnisse gleichsam als das Resultat eines ‚social contract' betrachtet werden können, d.h. als eine Rechtskonstruktion, in der jeder seine Ansprüche nach Maßgabe seiner individuellen Anfangsfestlegung befriedigt bekommt.[13] Die Marktallokation wird als ein Interaktionsverfahren beschrieben, über dessen Gesamtergebnis — im ‚Gleichgewicht', aufgeteilt in die vielen individuellen Einzelergebnisse — *nicht mehr verhandelt werden kann* (Ausschluß von ‚recontracting'). Im Gleichgewicht wird das System geschlossen. So wird der Markt *redistributionsneutral* gedeutet: *Gegen die — im Markt repräsentierte — Gesellschaft haben die Individuen keine weiteren Ansprüche als die, die sie nicht bereits selbst — durch ihre eigene, unverrückbare Entscheidung — vorgetragen haben.*

Dieses Ergebnis hat die bedeutsame moralische Implikation, daß niemand nachträglich Ansprüche (vornehmlich Redistributionsansprüche) an ‚den Markt' stellen kann, die er nicht vorher bereits geäußert hatte. Wir haben es mit einer impliziten Handlungsvorschrift oder -regel zu tun. Der Markt wird methodisch aus der Zone der Ungerechtigkeit, in den ihn die ältere Moralphilosphie gestellt hatte, befreit (vgl. *Gauthier* 1982). Doch beruht diese Neutralisation der Moral selber auf einer moralischen Implikation: Und zwar *auf der Zuschreibung von Verantwortlichkeit für das eigene Handeln der Individuen.*

Wir können, wenn wir dieser Argumentation folgen, deutlich sehen, daß die ‚Rationalität' eine moralische Konstruktion enthält. Sie unterstellt den Individuen, *daß sie, weil sie wissen, was sie tun, verantwortlich sein können.* Man findet diese Behauptung in den Perfektabilitätsannahmen wieder, d.h. in der Annahme ‚vollständiger Information'. Diese Annahme ist von H. A. Simon u.a.m. systematisch bezweifelt worden mit dem Hinweis auf die begrenzte Informationskapazität empirischer Entscheidungsvorgänge (*Simon* 1982; *Kahneman / Slovic / Tversky* 1982; *Tversky / Kahneman* 1990; *Slote* 1990). Ich halte diesen Einwand für berechtigt, wenn man empirische Verhaltensforschung betreiben möchte, möchte aber die rational choice-Konzeption an dieser Stelle gegen den Einwand verteidi-

---

12 So auch *A. Schütz:* Die Volkswirtschaftslehre „setzt als wissenschaftliches Postulat das Studium des Handelns der beobachteten ökonomischen Subjekte so an, *als ob* diese zwischen vorgegebenen problematischen Möglichkeiten zu wählen hätten" (*Schütz* 1971, 97).

13 Hier wird die Vertragslogik einer Tauschtheorie offensichtlich, die unterstellt, daß alle Verträge optimal sind, da es keinen Grund für ein frei und unabhängig agierendes Individuum gäbe, nicht-optimale Verträge einzugehen. Aus dem Faktum des geschlossenen Vertrages wird rigoros auf die Kompetenz der Vertragspartner rückgeschlossen. Empirisch sieht es allemal anders aus.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

gen: Die Unterstellung ‚vollständiger Information' ist in der rational choice-Theorie keine informations- oder kognitionstheoretische Behauptung. Es geht gar nicht um empirische Valenz. Sondern die Behauptung dient einzig und allein dazu, das Wahlhandlungssubjekt als einen entscheidungsfähigen Akteur zu konstruieren, *dem — in Analogie zur Rechts-Person — ‚die Tat zugerechnet' werden kann.*[14]

Die ‚Vollständigkeit' soll gar nicht in Informationsmaßen ermessen werden, sondern in der subjektiven Gewißheit, diese Entscheidung vor allen anderen richtig gefällt zu haben. Die Gewißheit hat an dieser Stelle eine systematische moralische Konnotation mit dem ‚Gewissen'. Ob das Individuum nun tatsächlich alles Ermessenmögliche getan hat, um sich dieser Gewißheit zu versichern, ist für die Theorie unerheblich. Es genügt vollkommen, ihm zu unterstellen, *daß es prinzipiell in der Lage ist, alle seine Möglichkeiten, ein Urteil zu finden, zu nutzen*[15] (in Abgrenzung zur Dezision, die keine rationale Rechtfertigung kennt [*Kersting* 1992, S. 24]).

Es kommt dabei einzig darauf an, eine beobachtete Handlung als *eindeutige und irreversible* Entscheidung auszuweisen. Denn wenn der Akteur bei einer Handlung beobachtet wird, wird ihm aus dem Umstand, daß er mehrere Alternativen gehabt haben kann, die geoffenbarte Präferenz als durch einen Wahlakt zustande gekommen zugerechnet. Daß diese Entscheidungszuweisung zugleich ‚rational' genannt wird, ist schlicht aus der Zuschreibungskonstruktion zu erklären. Man unterstellt, daß das Wahlhandlungssubjekt eine ‚vernünftige' Person ist[16]; im Zustand dieser ‚Vernunft' wird es nicht irgendeine Alternative wählen, sondern die unter den Umständen ‚beste'. Nur aufgrund dieser ‚Vernunft'-Prämisse kann die Verantwortlichkeit für die Entscheidung deklariert werden. Denn wenn das Wahlhandlungssubjekt lediglich irgendeine der möglichen Alternativen gewählt hätte (im Modus der Dezision), würde zwar auch eine eindeutige Entscheidung gefällt werden, aber die Irreversibilitätsbedingung verletzt.[17]

---

[14] Als ‚idealtypische' Konstruktion, wobei ‚Unwissen nicht vor Strafe schützt'. Es geht nicht darum, daß der Akteur rational handelt, sondern daß er hätte rational handeln können.

[15] *Harsanyi* z. B. konstruiert für die rational choice hypothetische ‚wahre Präferenzen', die ein Individuum haben würde, wenn es 1. über alle einschlägigen empirischen Kenntnisse verfügte, 2. seine Überlegungen mit der größtmöglichen Sorgfalt anstellte und 3. sich in einer für eine rationale Entscheidung günstigen psychischen Verfassung befände (*Harsanyi* 1982; vgl. auch *Birnbacher* 1990, S. 74). Der hypothetische Charakter dieser Konstruktion läßt sie nur normativ verwenden.

[16] *J. Elster* bezeichnet diese ‚Hyperrationalität' als eine Art von Irrationalität (*Elster* 1989, S. 31).

[17] Eine Dezision hat keinen ‚entscheidenden' Grund; folglich kann eine dezisionale Entscheidung nicht ‚besser' sein als jede andere der verworfenen Handlungsmöglichkeiten (*Kersting* 1992, S. 24). In diesem Sinne ist eine Dezision zwar eindeutig; es fehlt aber die rationale Begründung, die sicherstellt, daß alle anderen Alternativen irreversibel ausgeschlossen bleiben.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Eindeutigkeit der Entscheidung ließe sich auch durch einfache Dezision herstellen; erst die irreversibel-eindeutige Entscheidung aber ist ‚vernünftig', weil jetzt unterstellt werden darf, daß das Entscheidungssubjekt *dafür einstehen* wird, diese und keine andere Entscheidung getroffen zu haben. Wir können dies eine *‚self-commitment'-Bedingung* nennen oder eine deontische Komponente, ohne die die rational choice ihr Konsistenzkriterium verletzen würde.[18] Mit der Verantwortung für die ‚Vernünftigkeit' ihrer eigenen Entscheidung übernehmen die Individuen — zugeschriebenermaßen — eine Verpflichtung. ‚Verpflichtung' ist nur ein anderer Name für moralisches Sollen (*Tugendhat* 1993, S. 34) — das in diesem Fall nicht auf andere gerichtet ist, sondern auf das Individuum selbst im Sinne einer evidenten Kulturmaxime etwa der Art der ‚protestantischen Ethik', daß ‚jedermann sein persönliches Schicksal nüchtern zu meistern habe'.

Mit dem kulturell ausgeprägten Vernunftmuster, *daß die Individuen wissen können, was sie tun und dafür einstehen,* wird gleichzeitig unterstellt, daß dieses Wissen allgemeingültig ist, so daß niemand in Kooperationsverlegenheiten kommt, d. h. daß niemand darauf *angewiesen* ist, mitleidig oder solidarisch behandelt werden zu müssen.[19] Der Erfolg dieser Unterstellung beruht nicht darauf, daß sie aller emotionalen oder moralischen Einflüsse entledigt ist, sondern auf ihrer impliziten Zuschreibung von moralischer Verantwortung für das je eigne Tun und Lassen. Die theory of rational choice ist moralisch konfundiert bzw. genauer gesagt: *Ihre allgemeingültige normative Zuschreibung beruht auf dem Umstand, daß jedes Individuum nur dann regelmäßig rational handeln kann, wenn ihm die Konsequenzen seiner Handlungen — gleichgültig, wie sie tatsächlich ausgeführt worden sind — so zugerechnet werden können, daß es mögliche Enttäuschungen aus einem Mißlingen nicht sozial delegieren kann.*

3. Enttäuschung aus Mißlingen

Es wird normativ unterstellt, daß alle nach der Regel handeln, ‚handle rational'. Alle, die nach dieser Regel handeln, haben die Folgen ihres Tuns — ihre Erfolge wie ihre Enttäuschungen — mit sich selbst auszumachen. Für den Fall des Erfolges ist dies eine unproblematische Bedingung. Enttäuschungen dagegen werden nicht kompensiert. Die rational choice-Konzeption entwirft einen Individuum-Typus, dessen kognitive oder intellektuelle Kompetenz, die richtige Wahl zu treffen, gepaart ist mit einer psychischen Kompetenz, mögliche Enttäuschun-

---

[18] Das Konsistenzkriterium — so hat man sich in der Literatur scheinbar geeinigt — scheint die conditio sine qua non für die Rationalitätsdefinition zu sein (vgl. *Sugden* 1991; *Binmore,* 1987); auf Transitivität (*Fishburn,* 1988, S. 48) und den Ausschluß von Indifferenz läßt sich u. U. verzichten. Kritisch zum Konsistenzpostulat: *Nicolaides* 1990.

[19] Es bleibt in diesem Handlungskontext natürlich jedem freigestellt, mitleidig oder solidarisch handeln zu können; davon aber hängt das Markt-Handlungsergebnis systematisch nicht mehr ab.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

gen, die ihm seine rationale Entscheidung bereitet, zu absorbieren und mit sich selbst austragen zu können.

Dies ist keine Unterstellung, sondern eine systematische Implikation, wenn wir das dem rational actor zugeschriebene ‚self-commitment' betrachten. Wenn dem rationalen Individuum die volle Verantwortung für seine ‚Tat' zugerechnet wird, muß man — in einer rationalen Handlungstheorie — gleichzeitig unterstellen können, daß es kompetent ist, diese Verantwortung zu übernehmen. Erst dann kann man das rationale Procedere enttäuschungsneutral formulieren: Das Individuum hat *jede* Konsequenz seiner Entscheidung zu tragen — *die erwarteten positiven Konsequenzen ebenso wie die nicht ausschließbaren negativen.*

Wenn normalerweise der Moral vorgehalten wird, daß sie heroischer Menschen bedarf[20], haben wir es hier mit einem — bisher unbeachteten — ‚Heroismus' der Rationalität zu tun: *Dessen Leistung beruht in einem Aushalten von Enttäuschungen darüber, daß die individuellen empirischen Entscheidungen dem normativen Zuschreibungsstandard der Rationalität nicht genügen.*

Indem man die rationale Wahlhandlung — wie es gewöhnlich getan wird — allein und ausschließlich als einen kognitiven Prozeß unterstellt, wird sie als Alltagskompetenz eines jeden Individuums verwendet. Etwa in der Form: Wenn man fragt, welche von zwei gegebenen Alternativen er präferiere, wird jeder antworten können: diese oder jene. Dies erscheint als evident, d.h. keiner weiteren Begründung bedürftig, weshalb man von einem ‚Rationalitätsaxiom' redet (*Vilks* 1991, Kap. I.1.1.1.). Rationalität wird hier als allgemeinmenschliche Unterscheidungsleistung vorgeführt — als ein kognitiver Akt unmittelbarer Evidenz.[21]

Doch konfligiert die kognitive Selbstverständlichkeit, mit der der Wahlakt als ‚rational' jedem Individuum zugeschrieben wird, mit der Anforderung, seine Entscheidung ‚aushalten' zu können, wenn sie durch die Umstände der Situation zum Mißerfolg führt.

---

[20] Vgl. *Kliemt / Zimmerling* 1993. Wer ist schon bereit zu opfern in Hinblick darauf, für andere Vorbild zu sein, damit sie selber überlegen, ob sie nicht auch auf diese Art und Weise moralisch sein wollten, wenn diese pflichtethische Haltung ‚ausbeutbar' ist?

[21] Die Evidenz dieser Behauptung entsteht schlicht aus ihrer Frageform, die die Gegebenheit von Alternativen voraussetzt. Diese Gegebenheit ist in gewöhnlichen Handlungssituationen aber erst herzustellen; das Individuum muß die Situation erst als eine solche, in unterscheidbare Alternativen differenzierbare wahrnehmen. Gewöhnlich sind Entscheidungssituationen unscharf strukturiert. Aber selbst dann, wenn die Individuen mit solchen eindeutigen Entscheidungssituationen versorgt wären, ist es nicht sicher, daß sie rational entscheiden, d.h. eine der angebotenen Alternativen präferieren. Indifferenz ist nicht automatisch ausgeschlossen.

Aber auch selbst dann, wenn wir sie ausschließen können, ist es immer noch nicht sicher, daß die Individuen rational entscheiden, denn es kann ohne weiteres der Fall sein, daß das in konkreten Situationen betrachtete Individuum zwar die Alternativen unterscheidet, aber nicht diejenige ‚wählt', die es in der Situation für ‚besser' hält, sondern diejenige bestimmt, die es immer schon bestimmen wollte. Wir können eine affektive Bestimmung der — scheinbar ‚gewählten' — Alternative nicht ausschließen. Emotion matters.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Wieso kann bei rational actions von Enttäuschung die Rede sein? Wir haben im vorhergehenden Abschnitt dargestellt, daß eine normative rational choice-Theorie jede Handlung als rational betrachtet, gleichgültig wie sie tatsächlich ausgeführt wird. Normativ gesehen kann eine rational choice nicht ‚mißlingen'. Wenn wir aber die Umstände, Motive etc. einführen, die die von der normativen Theorie betrachteten tatsächlichen Handlungen begleiten, haben wir zu untersuchen, wieweit es mit der Fähigkeit der Akteure bestellt ist, mit nichteintreffenden rationalen Erwartungen umgehen zu können. Denn: Wenn Enttäuschungen systematisch zum Entscheidungsvorgang gehören, dann ist die kognitive Gewißheit der rational choice — wiederum systematisch — allein schon durch den Umstand in Frage gestellt, daß etwas, was nicht eintreffen und darum enttäuschen *kann,* kognitiv auch nicht gewiß sein kann. Das Erkennen und Wahrnehmen von Alternativen ist noch kein Grund für eine erfolgreiche Realisation. Folglich ist die Entscheidungskompetenz nicht hinreichend bestimmt, wenn sie allein als Kognitionsvorgang eingeführt wird.

Natürlich ist es für die normative Theorie leicht zu sagen, daß ein mit einer gewissen Eintrittswahrscheinlichkeit bedachter und rational gewählter Erwartungsnutzen auch dann weiterhin ‚als rational gewählt' zu betrachten ist, auch wenn das Ereignis nicht eintritt. Dieses normativ-logische Resultat aber ist eine Rationalitätszuschreibung, die nichts darüber besagt, wie das faktische Nicht-Eintreffen des Ereignisses *vom Akteur aufgefaßt wird.* Die Diskrepanz zwischen rationaler Wahl und empirisch möglichem Verfehlen des erwarteten Ereigniseintritts kann vom rational actor nur dann wie selbstverständlich zum rationalen Ergebnisbereich gezählt werden, wenn die normative Rationalitätszuschreibung dem empirischen Verhalten zufällig völlig entspricht.

Die Pointe dieser Darlegungen besteht darin, daß die normative Rationalitätszuschreibung nur für einen Teil der Handlung zutrifft: für die rational choice (under uncertainty), nicht aber für das finale Ergebnis. Da eine Handlung aber erst dann vollständig beschrieben ist, wenn das intendierte Ergebnis erlangt ist, ist die rational choice-Zuschreibung eine unvollkommene Handlungserklärung. Wir haben die Sache nun so zu betrachten: Wenn das — mit einer gewissen subjektiven Wahrscheinlichkeit des Eintritts — rational erwartete Ergebnis nicht eintritt, ist die Intention der Handlung 1. nicht realisiert worden bzw. die Handlung ist nicht erfolgreich gewesen, obwohl sie 2. rational erfolgte. Daraus folgt, daß die rational choice nicht automatisch mit erfolgreichem Handeln korreliert, mit der Konsequenz, daß die Akteure sich überlegen, wenn sie nicht erfolgreich waren, ob es sinnvoll ist, ihr Handeln weiterhin in terms of rational choice zu beschreiben.

Denn würden die Individuen aus ihrem Mißerfolg — ihr ‚Verwählen' — keine Schlüsse auf ihr künftiges Verhalten ziehen, hätten wir es hier mit einer quasi-stoischen Attitude zu tun, die den Individuen unterstellt, daß sie ‚leidenschaftslos' auch die negativen Ergebnisse ihrer rationalen Wahlhandlungen akzeptieren. Vor allem aber: Daß sie in der Lage sind, unberührt vom vorhergehenden Ausgang

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

der Handlung die nächste Handlung wieder rational angehen zu können. Hierbei wird eine ‚identitätsphilosophische' Unterstellung gemacht, die dahingehend lautet: Das Individuum bleibt konstant dasselbe, gleichgültig welche ‚Erfahrungen' es mit seiner Rationalität macht.[22]

Es scheint etwas vermessen zu sein, mit der ‚schlechten Erfahrung' von rational choices die rational choice selbst gleich in Frage zu stellen. Man könnte erst einmal entgegnen, daß man daraus ‚lernen' sollte, d. h. künftig die Situation ‚besser' einschätzen. Wie aber, wenn man nicht wissen kann, was es heißt, eine unbekannte Situation einzuschätzen? Auch lohnt es sich nicht zuzuraten, ‚bessere' subjektive Wahrscheinlichkeiten zu bilden. Wonach sollen sie gebildet werden als nach subjektiven Einschätzungen? Subjektive Einschätzungen zu verbessern, erbringt allemal nur wieder subjektive Einschätzungen, die — als Wahrscheinlichkeitsaussagen — logisch immer die Möglichkeit des Nichteintreffens enthalten. Wenn aber logisch die Möglichkeit des Nichteintreffens dessen, was man subjektiv in seinem Handeln entschieden hat, nicht zu umgehen ist, dann ist die Verantwortungszuschreibung für die eigene ‚Tat' kein unproblematischer Vorgang. Denn trotz rationaler Unterstellung, die Möglichkeit des Nichteintreffens ex ante gewußt zu haben, kann das Faktum des Nichteintreffens ex post dennoch Enttäuschung hervorrufen.

Diese These ist für das Verhältnis von Moral und Rationalität von Interesse, da wir es in der rational choice mit einer Akteurskonstruktion zu tun haben, die konstant kognitiv operiert und durch ihre Handlungen bzw. durch ihre Ergebnisse nicht emotional affiziert wird. *Wir haben es mit einer optimistischen Akteurskonstruktion zu tun, der z.B. eine ‚Leidens'-Semantik fehlt.*

So werden wir darauf aufmerksam, daß die Theoriekonstruktion der rational choice bestimmte Beweggründe menschlichen Handelns — mit Gründen — außer acht läßt, vornehmlich alle Beweggründe emotionaler Art. Diese Beweggründe sind ‚moral sentiments' im weitesten Sinne. Es sind Gründe, die der Moral einen Spielraum lassen, den die Ökonomie gar *nicht füllen will.*

Wenn wir der Moral den Spielraum zubilligen, die individuellen und sozialen Dilemmata zu lösen, die den Beteiligten in irgendeiner Form ‚Leid' zufügen, können wir das erst einmal nur in moralischer Terminologie tun. Hier ist eine Grenze des ‚ökonomischen Imperialismus' erreicht, die sich nur unzureichend in einem Opportunitätskostenkalkül abbilden läßt. Man kann ‚Leid' nicht ohne

---

[22] Man könnte versucht sein, von einer ‚anthropologischen Konstante' zu sprechen; ich halte eher dafür, daß die implizite Identitätsbehauptung methodische Gründe hat. Sie bezeugt ein Verbot der Selbstreferenz, d. h. der Rückwirkung des Handlungsergebnisses auf die Entscheidungskompetenz des Individuums. Wäre das der Fall, müßte man von Situation zu Situation mit verschiedenen Entscheidungskompetenzen desselben Individuums rechnen, was das Problem mit sich brächte, bei schlechter Entscheidungsverfassung sich für ‚unzurechnungsfähig' zu erklären und also die Verantwortung nicht zu übernehmen. Die ‚self-commitment'-Bedingung würde entfallen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

weiteres in ‚Kosten' übersetzen. Wenn ‚Kosten' als ‚Nutzenentgang' opportunitätskostentheoretisch interpretiert werden, ist damit nicht das Leid erfaßt, das die Moralphilosophie thematisiert. Das ‚Leiden' daran, etwas nicht bekommen zu haben und sich mit etwas geringer Bewertetem zufrieden zu geben, ist nicht *das Leid, das unverfügbar über einen kommt und von der Person dergestalt Besitz ergreifen kann, daß sie gar nicht mehr handlungsfähig ist, d.h. nicht mehr ‚rational' agieren kann.*

Diese Behauptung wird nicht sogleich von jedem geteilt werden. Man könnte einwenden, daß ökonomische Handlungen meistens von einer Art sind, in der selbst bei Enttäuschungen das Leid nicht Besitz von der Person ergreift. Doch macht man mit dieser Argumentation den Fehler, ‚ökonomische Handlungen' in der Art und Weise aufzufassen, wie wir sie hier gerade in Frage stellen. Es gibt keine ‚ökonomische Handlung', die so standardisiert ist, daß sie ‚Leid' ausschließen kann (wenn wir zugleich ausschließen, daß ‚ökonomische Handlung' nicht einfach normativ Leid ausschließt). Wir reden nicht von ‚Äpfeln' und ‚Birnen' und anderen mikroökonomischen Lehrbuchdemonstrationsgütern. Wenn die rational choice *auf alle möglichen* sozialen Aktionen ausgeweitet wird, wie wir es — durchaus legitim und nicht erst seit G. S. Becker — in dem Trend der ökonomischen Theorie vorfinden, der ein Hang zum ‚ökonomischen Imperialismus' nachgesagt wird, dann haben wir es nicht mit solchen Trivialitäten wie einer Enttäuschung darüber zu tun, daß man auf etwas zu erwerben verzichten muß, wenn man etwas anderes will. Es geht hier nicht um konsumpsychologische Erwägungen, sondern um die Enttäuschungen, die systematisch dann eintreten können, wenn man (rationale) Erwartungen über das Verhalten anderer anstellte, um daraus einen Nutzen zu ziehen. Wir haben es hierbei mit ‚sozialen Enttäuschungen' zu tun, die existentiell wirken können — weil sie 1. entweder größere Vorleistungen entwerten oder weil sie 2. einen erwarteten Status verletzen etc. Doch sind auch diese beispielhaften Andeutungen weniger wichtig als der Umstand, daß die rational choice in ihrer bayesianischen Erwartungsnutzenvariante systematisch auf eine Gewinn- / Verlust-Auszahlung spekuliert, die für die Akteure eine neutralisierende Komponente eingebaut hat, bei Verlust oder Nichteintreffen des erwarteten Ereignisses stoischer Natur zu bleiben.

Eben diese Implikation täuscht: Es gibt keine Verhaltensstandards, die das Eintreffen erwarteter Ereignisse äquivalent ihrem Nichteintreffen behandeln. Die Rationalität der Bayes-Form des Entscheidungsverhaltens erfordert aber gerade diese Attitüde. Folglich ist es zulässig — und zwar aus systematischen Gründen —, diesem Rationalitätstypus eine Inkompetenz im Umgang mit Mißerfolgen zuzusprechen: Der Umgang mit Mißerfolgen kann aus der logischen Struktur der rational choice heraus nicht thematisiert werden.

Zur Ergänzung sei darauf hingewiesen, daß der scheinbare Vorteil dieser Konstruktion, nicht-emotional auf Verluste zu reagieren, dann zum Nachteil dieser Konstruktion wird, wenn die Folgen eines Verlustes individuell nicht

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

verkraftet werden. In diesem Fall ist es zwecklos, darauf zu verweisen, daß der Aktand dies hätte wissen und einkalkulieren müssen. Da die Attitüde der rational choice eine normative Zuschreibung darstellt, keinen Ausdruck des empirischen Verhaltens oder Verhaltenkönnens, muß faktisch mit ‚Leid' auch in den Fällen gerechnet werden, in denen die Akteure mit wohlanständigem ‚rational belief' ihre Entscheidung fällten. *Sie erweisen sich dann als ‚schwächer', als die rationale Norm erlaubt.* Dieses Ergebnis kann eine ökonomische Handlungstheorie nicht ignorieren, weil es den Akteuren Anlaß gibt, sich künftig nach anderen Kriterien zu verhalten als denen, die die rational choice ihnen vorschreiben möchte. Und zwar dann aus der Erfahrung mit rationalen Handlungen.

‚Leid' in diesem Sinne ist kein verfügbares Objekt einer Wahl, kein ‚Gut'. Zudem würde auch niemand ‚Leid' wählen, höchstens Leidensunterschiede in der Absicht, weniger leiden zu müssen. Doch ist das ‚Leid' — das nicht nur die christliche Moralphilosophie zum Zentralpunkt moralischer Behandlung erkürt[23] — per definitionem von einer Art, *in der der Leidende über keine Handlungsmöglichkeiten mehr verfügt.* Daran zeigt sich eine basale Differenz zur rational choice, die ausschließlich als *Theorie der Handlungsverfügbarkeit* auftritt. Somit reicht es für unsere Argumentation, darauf hinzuweisen, daß die rational choice als Handlungsschema in unsicheren Situationen durchaus versagen kann[24], was nichts anderes bedeutet, als daß die rational choice nicht von vornherein identisch ist mit der Kompetenz, die Handlung auch zu realisieren.

Moralisch bedeutsam wird dieses Versagen allerdings erst dann, wenn ihm ein ‚Leiden' korrespondiert, das aus der Enttäuschung des Versagens entstand. Das ist nicht immer der Fall; die meisten ökonomischen Handlungsversagen werden vielleicht Ärger und Wut erzeugen, nicht aber das gemeinte ‚Leiden'. Doch gibt es natürlich hinreichend Fälle, die den erwünschten Erfolg der rational action so eng mit dem Selbstbild des Individuums verbinden, daß ein Versagen kritische Auswirkungen hat.

Das Problem des ‚Selbst' wird für die rational choice dort bedeutsam, wo die normative Unterstellung der ‚Identität' des Wahlhandlungssubjektes problematisch wird, und damit seine overall-Kompetenz, rational zu handeln. *Am ‚Leid', das handlungsunfähig macht, ist die Grenze der rational choice erreicht, und wir betreten die Abteilung der Moral, die sich um die Person-Identität sorgt.*[25] Man muß natürlich hinzufügen: Nur dort ist die Grenze der rational choice

---

23 Vgl. *Schopenhauer* oder — um einen modernen Philosophen zu nennen — *Richard Rorty* 1989 (bes. die Schlußkapitel).

24 Vgl. *Ellsberg* 1990. Aber auch in sogenannten ‚sicheren' Situationen (unter ‚Gewißheit'). Genau betrachtet kann es keine ‚sicheren' Situationen geben, wenn wir bedenken, daß jede kognitive Eindeutigkeit ein Interpretat des Handlungssubjektes ist. Wir können dann lediglich behaupten, daß das Subjekt meint, sicher gewissse Alternativen unterscheiden und zwischen ihnen entscheiden zu können (vgl. *Rector* 1991).

25 *E. Tugendhat* ist der Meinung, daß das die Zentralabteilung der Moral sei (*Tugendhat* 1993).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

erreicht. Wir analysieren die Bedingungen und Möglichkeiten des rational choice-Ansatzes, nicht seine Infragestellung.

Der ‚logische Ort' der Moral in der Ökonomie befindet sich an der Stelle, in der die — die Ökonomie fundierende — rationale Handlungstheorie das Scheitern des Wahlhandlungsaktes nicht thematisieren kann. Bzw. genauer: Die rational choice-Theorie kann Mißerfolge nur als Chance neuer rationaler Handlungsoptionen ‚verstehen'. Damit aber wird das individuelle Scheitern an der Marktrealität unmittelbar in terms of ‚sunk costs' übersetzt, die den Akteuren erlauben, in der nächsten Handlung denselben Rationalitätsoptimismus zum Einsatz zu bringen, mit dem sie vorher gescheitert waren. Die Möglichkeit, daß die Individuen ihr Scheitern als Mißerfolg interpretieren, wird methodisch ausgeschlossen.[26] Damit wird auch ausgeschlossen, daß sie aus Mißerfolgen ‚handlungsunfähig' werden (was die Unfähigkeit zu rational actions einschlösse, also ausgeschlossen werden muß).

An dieser Stelle muß ein Einwand erörtert werden. Manche Ökonomen werden die vorgetragene Interpretation zurückweisen, weil sie nicht sehen, wie rational choice-Handlungen und Mißerfolge / Enttäuschungen zusammenhängen können. Auch auf die Gefahr einer Wiederholung ist dieser Punkt zu klären wichtig. Nur wenn die rational choice eindeutig diejenige Alternative als ‚beste' auswählt, die unmittelbar für den Akteur zur Verfügung steht, gibt es eine eindeutige Intention / Ergebnis-Relation. Für die rational choices nach dem Modus der bayesianischen expected utility ist diese Relation durch ein Eintrittswahrscheinlichkeitsmaß gebrochen. So wie dem bayesianischen Akteur zwei Ergebnisse möglich sind: 1. daß das erwartete Ereignis eintritt, 2. daß es nicht eintritt, so ist das negative Ergebnis wiederum zweifach zu bewerten – a) kann der Akteur der ‚rational belief' anhängen, daß das negative Ergebnis (2) konsequenterweise ein rationales Ergebnis ist. Oder aber er kann b) das negative Ergebnis als ‚Verwählen' interpretieren, d.h. als einen Mißerfolg, der nicht mit der ursprünglichen ‚rational belief' kongruiert.

Position a) wird von den Akteuren vertreten, die ihre Handlungen normativ rational interpretieren. Position b) wird von den Akteuren vertreten, die in Zweifel geraten, ob ihre Handlung rational genannt werden kann. *Das Problem, das mit diesen Unterscheidungen gezeigt werden soll, besteht schlicht darin, daß es keine rationale Entscheidung darüber geben kann, welche Position einzunehmen ist.* Beide Positionen sind sinnvolle Interpretationen des Nicht-Eintritts des rational

[26] Oder durch ‚Lernen' thematisiert. ‚Lernen' aber ist nur eine andere Variante des prinzipiell unterstellten Handlungsoptimismus. Wenn man sagt, daß ein Akteur ‚aus Fehlern lernt', werden die Mißerfolge in Informationen verwandelt für neue Erfolge. Abgesehen von den Problemen der informationsökonomischen Unterstellungen, das ‚Wissen' aus einer vergangenen Situation auf neue (andere) Situationen einfach übertragen zu können, geht es im obigen Kontext um jene Fälle, in denen die rationalen Erwartungen enttäuscht werden, und zwar so, daß die Enttäuschung auf die Bildung von rationalen Erwartungen übergreift.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

erwarteten Ergebnisses. Man könnte zwar sagen, daß jeder bayesianische Akteur hätte wissen können, daß seine subjektive Eintrittswahrscheinlichkeitsbehauptung nicht effektiv werden kann (weil sie mit einem Wahrscheinlichkeitsmaß operiert), dann aber lassen sich nicht die Zweifel ausräumen, wozu die rationale Theorie der expected utility nützlich sei. Die Antwort lautet natürlich: Sie ist nützlich, weil sie mit einer höheren Wahrscheinlichkeit den Eintritt des erwarteten Ereignisses beschreibt. Doch ist auch dies nur eine hypothetische bzw. eine statistische Antwort. Niemand weiß, ob die höhere Wahrscheinlichkeit im Wahlhandlungsakt realisiert wird. Man kann dabei nur Aussagen über längere Sequenzen derselben Wahlhandlung machen. Da die meisten rational choices aber nur einmalige Wahlhandlungen sind (*Simon* 1983, S. 42), gibt es keine Dauer, in der sich die relativ höhere Eintretenswahrscheinlichkeit zeigen könnte. Da es keine Wahrscheinlichkeitsverteilung darüber gibt, ob in dem jeweils vorliegenden Einzelfall die positiven Ergebniserwartungen eher eintreffen als die negativen, bleibt der Ausgang kontingent.[27]

Wenn aber bayesianische rational choice für den Einzelfall kontingent bleibt — weil wir nur selten identische Wahlhandlungen in identischen Situationen vorfinden —, und wenn wir mit dem Nicht-Eintreten der erwarteten Ergebnisse zu rechnen haben, dann kann zwar die normative Interpretation der Rationalität weiterhin behaupten, daß dies rational sei; die Akteure selbst werden es anders interpretieren. Wenn sie aber die Mißerfolge als ‚rationales Verwählen' interpretieren, entsteht eine Spannung zwischen ökonomischer-theoretischer Zuschreibung der Rationalität und einer davon getrennten Handlungsinterpretation der Akteure selbst. Diese Spannung kann die theory of rational choice nicht thematisieren.[28]

An diesem Punkt können wir nicht mehr von einer ‚Verantwortlichkeit des Individuum, für sich selbst' reden, da der Mißerfolg gerade die zugeschriebene

---

27 Wenn wir zudem behaupten, daß jede Entscheidung verschieden ist, da es keine völlig identischen Situationen gibt, ist eine bayesianische rational choice im Einzelfall eine kontingente Entscheidung, die eintreffen kann oder nicht, und zwar unabhängig von den subjektiven Eintrittswahrscheinlichkeitsvermutungen.

28 Handlungsunfähige Individuen sind eindeutig anderer Konstitution als handlungsfähige Individuen. Wenn aber durch die Selbst-Bewertung der Handlung als Mißerfolg die ‚Identität' des Individuums in Frage gestellt werden kann, kann die normative Zuschreibung von Rationalität nicht mehr umstandlos in Anspruch genommen werden. Bzw. sie kann nur unter der besonderen Restriktion weiter aufrechterhalten werden, daß die durch Enttäuschung am Mißerfolg handlungsunfähig gewordenen Individuen diesen Zustand eben selber zu verantworten haben. Der Mißerfolg gehört in den Konsequenzbereich ihrer Handlungsrationalität, weshalb sie ihn, da sie ihn als eines der möglichen Handlungsergebnisse erwarten mußten, ‚vernünftigerweise' akzeptieren müssen. Ein ‚Mißerfolg', den ich ‚vernünftigerweise' erwarten muß, ist aber kein Mißerfolg im pathischen Sinne, sondern ein normales Handlungsergebnis. Folglich ist der ‚Mißerfolg' semantisch für die rational chocie-Theorie gar nicht definierbar, obwohl er von den Akteuren sinnlich-pathisch erlebt wird. Die rational choices produzieren gleichsam ‚emotional outputs' (als affektive Externalitäten), die von der Theorie nicht formulierbar sind, weil der Gebrauch der Vernunft-Semantik ihre Existenz ausschließt.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Selbst-Verantwortung in Frage stellt. Die der rational choice-Konstruktion implizite Moral der Selbst-Verantwortung kann nur dann als moralische Implikation weiterhin Geltung beanspruchen, *wenn sie die Akteure handlungsfähig erhält.* Die ‚vernünftige' Zurechnung der Verantwortung des Individuums für seine Handlungen (und deren Folgen: Erfolg oder Mißerfolg) verliert ihren Sinn, wenn das Ergebnis — wie im Enttäuschungsfall möglich — zur Handlungsunfähigkeit führt. Das Problem der möglichen Enttäuschung an der Stelle, wenn das ‚rational design' der Handlung mit anderen Ergebnissen konfrontiert wird, die der Akteur nicht erwartet hatte, kann nicht voluntaristisch-vernünftig gelöst werden, indem der Betreffende ‚das nächste Mal' bessere Entscheidungen ausführt. *Wenn die rational choice-Konzeption behauptet, in der Präferenzordnung den ‚emotional input' sortiert und einer rationalen Evaluation übereignet zu haben, kann sie hingegen den ‚emotional output' nicht regulieren.*

Der emotionale Tensor des Wunsches übergreift den rational choice-Prozeß und kann sich in der Enttäuschung über einen Mißerfolg wieder zeigen. Das heißt nichts anderes, als daß die Bändigung der Leidenschaften, die der rational choice-Prozeß versprach, nur kontingent gelingt, nämlich immer nur dann, wenn das tatsächliche Verhalten in terms of normative rational choice interpretierbar ist — und *von den Akteuren in dieser Interpretation auch als sinnvoll verstanden wird.*

4. Beobachter- und Akteursverhalten

Somit haben wir es mit dem eigentümlichen Umstand zu tun, die normative Rationalität nur in optimistischen Zusammenhängen brauchbar zu finden, nicht aber in Versagenstatbeständen. Es geht um das Problem der rational choice-Konzeption, den Individuen kognitive Rationalitätskompetenz zu unterstellen, die sie im Falle des Mißerfolges nicht durchhalten können. Denn ähnlich wie den bayesianischen Akteuren zwei Grundtypen des Risikoverhaltens unterstellt werden (risikoaversiv / risikofreudig), müssen den rational actors, wie wir es im vorheringen Abschnitt bereits dargelegt hatten, zwei Typen der Mißerfolgsverarbeitung zugerechnet werden: Entweder halten sie es für ein normales Risiko ihrer rationalen Wahlhandlungen (I) oder aber für eine Inkompetenz (II). Im letzteren Fall ist nicht mehr garantiert, daß ihnen rationales Verhalten im unterstellten kognitiven Sinne zugeschrieben werden kann. Für enttäuschte Akteure müssen wir mit der Zunahme von Indifferenz (Unentscheidbarkeit) rechnen[29] oder allgemeiner mit „imperfect decisions" (*Heiner* 1988). Zumindest können wir es nicht ausschließen, daß sie, anstatt jeweils wieder neu rational zu entscheiden, sich an ältere, erfolgreiche Erfahrungen anlehnen, um in neuen Situationen dieser Erfahrung wie einer Regel zu folgen.[30]

---

[29] *Sugden* 1991; *Elster* 1990. Oder auch in der Form der „nontransitive indifference" (vgl. dazu: *Fishburn* 1988, S. 40 ff.).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Spannung zwischen normativer Rationalitätszumutung und empirisch nicht-rationalem Entscheiden — die in der neuen mikroökonomischen Literatur an der Zunahme der Erörterungen von ‚Paradoxien' und Entscheidungsanomalien sichtbar wird (vgl. *Sugden* 1991; *Frey / Eichenberger* 1989; *Heap* 1988, Kap. 2; *Nicolaides* 1990; *Davidson* 1990 (II); *Slote* 1990) — muß vom rational actor individuell verarbeitet werden. Viele Wahlsituationen stellen sich für die empirischen Subjekte als indifferent dar. Wenn sie dann irgendeine der Alternativen wählen, haben sie nicht rational gewählt, weil sie nicht ‚die beste' ausgesucht haben. Zwar können wir sagen, daß die Individuen im Indifferenzfall eben mehrere beste Alternativen haben, zwischen denen sie kontingent ‚wählen' können, ohne das Rationalitätskriterium zu verletzen. Wir haben dann formal das Rationalitätsprinzip bei Indifferenz bewahrt, aber ein anderes Problem eingehandelt: eine ungenügende Differenzierungsfähigkeit bei der Herstellung des Wahlbereiches.

Formal ist es richtig, alle indifferenten Alternativen gleichwertig als ‚eine beste Alternative' zu bezeichnen. Doch ist damit die rationale Wahl in zwei Wahlprocedere aufgespalten: in die Wahl — genauer: die Unterscheidung — von n und a, bevor innerhalb von n eine Dezision getroffen wird (n ist nur ein Name für eine Reihe von Alternativen, z.B. von b, c, d, e, f. In der Vor-Wahl werden b, c, d, e, f als besser als a deklariert, womit die Bedingungen einer rationalen Wahl erfüllt sind: $a < n$). Die Entscheidung zwischen b, c, d, e und zum Beispiel f wird zwar *als Entscheidung* übriggelassen, aber *nicht als rationale Wahl.* Wenn das Individuum sich zwischen den indifferenten Alternativen b, c, d, e, f entscheidet, trifft es demnach jedesmal eine rationale Wahl, und zwar deshalb, weil die gesamte Gruppe — als n zusammengefaßt — bereits rational vorgewählt ist.

Scheinbar kann das Individuum zwischen a und n ($n = b, c, d, e, f$) ebenso klar unterscheiden und damit rational wählen wie zwischen a und zum Beispiel b. Doch ist dies eine Kompetenzunterstellung, die das *Unvermögen,* zwischen b, c, d, e, f hinreichend Alternativen unterscheiden zu können, in eine Fähigkeit umschreibt, b, c, d, e, f als n von a unterscheiden zu können. Deutlicher wird diese Unfähigkeit, wenn $a > n$ gilt. Dann zeigt sich, daß das Individuum zwar *eine* Unterscheidung machen kann, aber nur deshalb, weil es nicht in der Lage ist, alle anderen möglichen Alternativen präferentiell zu ordnen. Formell ist die

---

30 Vgl. dazu: *Heiner* 1988; *Heiner* 1983; *Selten* 1990; *Klopstech / Selten* 1984.
Regelgebundenes Verhalten besteht in der Antizipation von Unsicherheitslagen, auf die die Individuen nicht mehr mit subjektiver Wahrscheinlichkeitsbildung reagieren, sondern mit ‚Faustregeln' oder ‚Erfahrungswerten'. Anders formuliert: In der Rolle bayesianischer Akteure fühlen sie sich unwohl; risikoavers gestimmt, wollen sie ‚keine Risiken eingehen', von denen sie eher Mißerfolge als Erfolge erwarten.
Regelorientiertes Verhalten dieser Art ist aber nicht mehr uneingeschränkt rational zu nennen, da es nicht mehr die aktuelle Handlungsalternativensituation optimiert, sondern vergangene Optimierungen regelmäßig (schematisch) auf gegenwärtige, völlig andere Situationen überträgt. Regelorientiertes Verhalten ist weder irrationales noch nichtrationales, sondern ein ‚eingeschränkt rationales Verhalten', das auf Optimierung verzichten muß.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Unterscheidung von *a* und *n* weiterhin eine Basis für rationales Wählen. Die Übersetzung von *b, c, d, e, f* in *n* simuliert zwar eine rationale Reduktion einer Alternativenmenge in eine einzige Alternative, ohne aber das Unvermögen zu reflektieren, innerhalb von *n* Unterscheidungen herzustellen (um danach dennoch *irgendeine* der von *n* umschriebenen Alternativen zu ‚wählen').

Ich plädiere dafür, die Indifferenz als eine Faktum individuellen Handelns anzuerkennen, es aber nicht artifiziell in eine rational choice zu übersetzen. Im Indifferenzfall kommt den Individuen die Situation nicht so strukturiert vor, daß sie sie rational lösen können. Entscheiden sie sich dann aber für irgendeine Alternative, wird ihnen diese Entscheidung zwar als rationale Entscheidung zugerechnet — d.h. als ihre optimale Lösung, die sie zu verantworten haben —, aber sie können sich mit der Verantwortungszuschreibung nicht einverstanden erklären, da sie wissen, daß sie im zweiten Schritt innerhalb der Indifferenzmenge nicht-rational entschieden haben, sondern dezisional. Zwischen der (externen) normativen Rationalitätszuschreibung und ihrem eigenen Handlungswissen klafft eine Lücke, die als diffuse, emotionale Spannung erlebt werden kann (über die man nicht sinnvoll kommunizieren kann, vor allem häufig nicht mit Ökonomen, um nicht als ‚irrational' zu gelten).

Zumindest existieren parallel zwei Interpretationssemantiken: Wir können sie a) eine Beobachtersemantik nennen und b) eine Akteurssemantik. Eine solche Unterscheidung kommt in der Theorie rationalen Handelns nicht vor. Was der (ökonomische-theoretische) Beobachter ‚sieht', schreibt er dem Akteur als rational choice zu. Der Akteur selbst beschreibt sein Handeln nur dann als ‚rational', wenn er die Semantik der rational choice-Theorie gelernt hat. Wenn der Akteur sein eigenes Handeln anders beschreibt, muß das den Beobachter solange nicht interessieren, solange den Akteuren die Verantwortung für ihr Tun und Lassen zugeschrieben werden kann.

Doch erodiert die der rational choice implizite Verantwortungsmoral, wenn die Akteure selbst die Verantwortungszuschreibung nicht mehr akzeptieren, d.h. wenn sie sich in der Semantik der Beobachter-Rationalität nicht mehr zureichend beschrieben finden. Sie beginnen nicht nur, andere Handlungserklärungen zu kommunizieren, sondern sie verhalten sich auch nach diesen anderen Erklärungen (die ihnen ‚richtiger' vorkommen).

Es geht mir hierbei um folgendes: Wenn die meisten Entscheidungen eher emotionaler Art sind, gleichzeitig aber (ökonomische) Sprachnormen herrschen, in denen die individuellen Handlungen nur als rationale Wahlakte kommuniziert werden, entsteht die eigenartige *Anomalie, daß die Individuen ihr Handeln nicht mehr sinnvoll beschreiben können.* Sie wissen zwar ungefähr, was sie tun, haben aber für viele ihrer tatsächlichen Handlungsabläufe keine ökonomische Sprache, so daß sie ihr Tun als etwas anderes (nämlich als ‚rational action') beschreiben müssen, und zwar auch dann, wenn darin ihr tatsächliches Handeln nicht mehr vorkommt. Unter diesen Bedingungen werden die eigenen Taten so erzählt, als

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

ob sie der Norm (der Rationalität) entsprechen. Dadurch tritt der eigentümliche Effekt ein, daß man tun kann, was man will, und es ist immer richtig, wenn es ‚rational nacherzählt' wird.

Das aber hat — wenn wir die subjektive Seite des Problems stärker betonen — Konsequenzen für das Entscheidungshandeln selbst: Egal, was man wählt, es ist immer rational, wenn man überzeugt ist, daß die Wahl die beste ist (‚rational belief').[31] Das gilt für alle möglichen Handlungen, auch für die, die für andere offensichtlich unsinnig oder ‚irrational' erscheinen. Wird jemand darauf hingewiesen, daß es doch offensichtlich besser sei, statt der offensichtlich ‚unsinnigen' eine andere Entscheidung zu treffen, kann derjenige jederzeit darauf verweisen, daß er im Moment der Entscheidung das Gefühl hatte, diese und keine andere Entscheidung sei die richtige. Er wird sein affektives Muster als kognitives verteidigen. Diese Gewißheit reicht aber völlig aus für die Zuschreibung normativer Rationalität, wenn wir das subjektive Kriterium ernst nehmen.

Zugleich aber zeigt es, daß sich die Rationalität ohne weiteres auf einen Affekt begründen läßt. Wir können die individuelle Entscheidung nicht danach analysieren, ob jemand hinreichende Kenntnisse über seine Alternativen hatte und sie auch alle abgewogen hat. Zwar erwartet man irgendwie von einem Rationalitätskriterium, daß es den Bedingungen perfekter Information genügen sollte[32], aber wir haben gezeigt, *daß es allein auf den Kontingenzausschluß ankommt, d.h. auf die Tatsache, daß eine Alternative ausgewählt wird und daß das Entscheidungssubjekt auf diese Wahl festgelegt werden kann.*

Wir kommen nicht umhin, folgenden Tatbestand anzuerkennen: daß die rational choice nichts anderes ist als der Vorgang, irgendeine Alternative auszuwählen, die insofern als ‚rational gewählt' bezeichnet wird, als nur diese Alternative und keine andere *subjektiv für relevant* erachtet wird (*Fusfeld* 1992). Über die Umstände der Relevanz kann empirisch nichts Verallgemeinerbares gesagt werden.[33]

---

[31] „Wenn jemand, der rational ist, glaubt, daß er etwas glaubt, so hat er die Disposition, sich auf eine Weise zu verhalten, die ein optimaler Beweis dafür ist, daß er es (dieses Etwas) glaubt. Das ist einfach zu begründen. Wenn jemand rational ist und glaubt, daß er glaubt, daß *p*, dann wird er *p* sowohl seinen praktischen als auch seinen theoretischen Schlüssen als Prämisse zugrunde legen. In diesem Fall wird er sich nach der Annahme, daß *p*, richten und deshalb so handeln, als wenn er glaubte, daß *p*" (*Shoemaker, S.:* Rationalität und Selbstbewußtsein, in: Kienzle, B. / Pape, H. (Hrsg.): Dimension des Selbst, Frankfurt / M. 1991, S. 19 ff., hier: S. 31).

[32] was der ‚bounded rationality' Simons Anlaß gibt, die empirische Möglichkeit perfektiver Information zu bezweifeln und statt dessen ein ‚satisficing-behavior' vorzuschlagen. Doch bleibt Simons Lösung in seiner empiristischen Neigung stecken; er sieht nicht, daß es der normativen Rationalität nicht um empirisch perfekte Bedingungen geht, sondern um logisch kohärente. Die Unterstellung der logischen Kohärenz als Rationalität des Ökonomischen und Alltagshandelns ist historisch kulturell bedingt.

[33] Jede andere Behauptung führte die normative Rationalität in den Teufelskreis präziser (situativer) Bedingungs- und Nebenbedingungsfestlegung, die für jede individuelle Entscheidung gesondert angeben können lassen müßten, welche nun galten oder nicht, um daran zu vergleichen, ob das wählende Subjekt alle Bedingungen erfüllt hat

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Mehr kann die Rationalitätsannahme in der Ökonomie nicht bedeuten. Ob das betreffende Individuum die Relevanz der Alternative einfach so bestimmt hat oder durch kluges Abwägen gegen alle anderen Alternativen (welche werden überhaupt zum relevanten Alternativenraum hinzugerechnet?[34]), bleibt nicht nur unbestimmbar, sondern ist für die Rationalitätszumutung selber irrelevant.

Der kommt es wesentlich darauf an, *das betreffende Individuum fortan so behandeln zu können, als ob es alle Alternativen gegeneinander abgewogen hat:* Jeder individuelle Zustand, sei es ein affektiver oder ein vernünftiger, wird somit rational transformiert. Und zwar mit der Folge, daß das economic design damit rechnet, daß alle Individuen wissen, daß sie so beobachtet und behandelt werden, als wüßten sie um alle Umstände und Folgen ihrer Entscheidungen.

Es ist einer Sozialtheorie — wie der Ökonomie — allerdings zuzumuten, die Bedingungen der Handlungsverfassung der Individuen zu analysieren, statt sie lediglich normativ zuzuschreiben. Das würde bedeuten, die Ökonomie von ihrer impliziten Moral der ‚Verantwortlichkeit für sich selbst' zu emanzipieren. Es ginge dann nicht mehr darum, den Individuen die Verantwortung für ihre Entscheidungen zuzuschreiben, sondern sie selbst darüber entscheiden zu lassen, inwieweit sie sie zu tragen bereit sind. Es muß deutlich werden, daß ‚rational zu handeln' nicht nur heißt, kluge Unterscheidungen in Hinblick auf eine subjektive Ernennung der ‚besten' Alternative zu machen, sondern mögliche Enttäuschungen auf sich zu nehmen. Wer das ausdrücklich weiß, wird anders handeln, d.h. stärker auf Versicherungskomponenten achten, auf den ‚common sense', der herrscht — in summa: regel- oder kooperationsorientierter handeln.[35] Schlicht bedeutet dies, daß individuelle Entscheidungen sehr viel stärker öffentlich erörtert werden — ohne daß damit bereits ein Öffentlichkeitsgrad benannt ist —, bevor entschieden wird oder daß die individuellen Entscheidungen sich stärker an das halten, was in der Öffentlichkeit gemeinhin als gültig angenommen wird. Regelorientiertes Handeln muß nicht allein und ausschließlich als individueller Rückgriff auf positive Handlungserfahrungen interpretiert werden (vgl. *Heiner* 1983), son-

---

(oder nicht). Erst dann könnte man — aus der Sicht der Theorie — dem jeweiligen Entscheidungssubjekt nachweisen, daß es bestimmte Bedingungen gar nicht erfaßt und mit bedacht hätte, um ihm einen Fehler in der rationalen Wahl nachzuweisen. Können wir das nicht — und wir können es nicht zuverlässig und ambivalenzausschließend —, dann können wir auch nicht behaupten, ein Verhalten danach abprüfen zu können, ob es rational war oder nicht. Das ‚interne' Rationalitätsprocedere ist nicht theoriefähig (zumal die Ökonomie dann von einer — bestimmten — Psychologie oder Neurophysiologie abhängig würde, was selber wieder methodische Probleme mit sich brächte: Warum gerade diese Psychologie und warum keine andere? etc.).

[34] Zur Problematik der Bedeutungszuweisung und des ‚Verstehens' von Alternativen vgl. verschiedene Aufsätze in: *Lavoie*, D. (ed.): Economics and Hermeneutics, London / New York, 1991.

[35] An dieser Stelle ist der Übergang zur soziologischen Theorie ‚sozialer Normen', deren Erklärungsvalenz möglicherweise wieder steigt (vgl. *Elster* 1990, S. 45 ff.; vgl. aber auch *Coleman* 1990; *Taylor* 1990 und *Hardin* 1990).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

dern kann ebenso eine kollektive Richtung einnehmen: daß man sich stärker an institutionellen Regeln orientiert oder schlicht daran, wie sich andere in der Situation verhalten (‚Imitation').[36]

Da die Wahl von Regeln weder notwendigerweise rational noch selbst wiederum durch Regelorientierung erfolgen kann — das Problem der Regelorientierung läßt sich in der Frage nach der Regel, an der man sich bei der Wahl von Regeln orientieren soll, als unendlicher Regreß formulieren —, ist es möglicherweise sinnvoll, anstatt wieder das normative Zuschreibungsgeschäft der Rationalität zu beginnen, die Kommunikationsstrukturen zu untersuchen, innerhalb derer Regeln gebildet werden.

5. Die Funktion von Moral in der Ökonomie

Wir hatten oben bereits erwähnt, daß das Rationalitätskonzept eher ein Bildungskonzept ist[37] als ein empirisch gehaltvolles Modell (was nicht ausschließt, daß das normative Schema der rational choice historisch so eingeübt ist, daß es als ein empirisches Schema gelten kann; wir können aber empirisch nicht feststellen, ob das so ist [*Davidson* 1990 (I)]). Wir hatten weiterhin erörtert, daß die rational choice-Konzeption implizit eine moralische Vorstellung von der Verantwortung der Individuen für ihre Handlungen enthält. Das ist zwar eine ungewöhnliche Formulierung des ‚methodologischen Individualismus', sachlich aber zutreffend. Die Individuen werden unter der Perspektive normativer Rationalität so betrachtet, als ob sie einen social contract geschlossen hätten, der ihren Autonomiestatus geklärt hat. Darin wird die Handlungsrationalität zur Norm erhoben, der sich niemand entziehen kann. Natürlich kann sich faktisch jeder dieser Norm entziehen, *aber der social contract bestimmt, daß er weiterhin so behandelt wird, als unterstünde er der Norm.*

Wir hatten weiterhin gesagt, daß es gälte, die implizite Moral der normativen Rationalität explizit zu machen. Dazu gehört erstens, daß die Verantwortung, die jedem Individuum zugeschrieben wird, von den Individuen selbst zu bewerten ist, ob sie sie auf sich nehmen können, zweitens aber, das Verhältnis von Moral und Emotion näher zu beleuchten.

Zur Verantwortung hatten wir bereits gesagt, daß sie eine Art umgekehrten Heroismus fordere. Die Individuen hätten die Spannung zwischen rationaler Zuschreibung und tatsächlicher, oft emotionaler Wahl auszuhalten. Einen mögli-

---

36 *Schnabl* 1987. Damit ist aber das ‚Unabhängigkeitsaxiom' der rational choice Theorie irrelevant geworden.

37 In seiner „Philosophy of Intransitive Preference" schreibt Paul Anand: „I do not even say that expected utility should never be used (it is of considerable pedagogic value), but I do believe we are in the midst of rethinking our ideas of rationality" (*Anand* 1993, 345).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

chen Defekt dieser Spannung hatten wir bereits erwähnt: Sie ‚verstehen' möglicherweise ihr Handeln nicht mehr.

Die Individuen geraten in eine Ambivalenz: Interpretieren sie ihr Handeln rational, geraten ihnen die tatsächlich handlungsleitenden Beweggründe, vornehmlich die emotionalen, aus dem Auge. Interpretieren sie dagegen ihr Handeln mit anderen, nicht-ökonomischen und nicht-rationalen Erklärungen, gerät ihnen ihre ökonomische Intention aus dem Auge. Sie handeln dann irgendwie, entlastet von der (impliziten moralischen) Anforderung der Rationalität, sich für die Folgen ihres Handelns verantwortlich zu zeigen. Es gibt dann keinen Grund mehr, nicht auch von affektiven oder voluntativen Unterscheidungen zu reden: das *will* ich und das nicht.[38]

Da wir den Ergebnissen der rational choices nicht ansehen können, ob sie eine rationale oder affektive Basis haben, können wir

1. feststellen, daß die ökonomische Theorie der Rationalität verträglich ist mit einer emotionalen oder affektiven Bestimmung der Entscheidung, auch wenn sie als ‚rational' zugeschrieben wird.
2. können wir feststellen, daß diese Unentscheidbarkeit zwischen kognitiver und affektiver Rationalität Konsequenzen hat für die normative Bestimmung der Rationalität: *Nur für die kognitive Version können wir die implizite Moral der Verantwortlichkeit für die eigenen Entscheidungen aufrechterhalten. Für die affektive Version muß das bestritten werden.*

Wenn die Entscheidung auf einer affektiven Bestimmung beruht, können wir nicht mehr unterstellen, daß die Individuen ‚so vernünftig' sind, zu wissen, was sie tun. Folglich können wir sie auch nicht mehr ‚verantwortlich' nennen für ihre Entscheidung. Wenn wir sie aber nicht mehr — kraft normativer Rationalitätszuschreibung — der impliziten Moral autonomer Individuen (die in Analogie zur rechtsfähigen Person entstand) unterstellen können, verliert die normative Rationalität ihre Legitimation, und *wir müssen neue moralische Institutionen erfinden, die auch die affektiven Entscheidungen verantwortbar machen. Das ist die methodische Anforderung an die Moral in der Ökonomie.*

Die Anforderung besteht nicht aus moralphilosphischen Anträgen an die Ökonomie, ist kein externer Anspruch (den man gegebenenfalls abweisen kann), sondern ist ein Anspruch, der aus dem Defekt der Selbstthematisierung der rational choice entsteht, sich ihrer potentiell emotionalen Entscheidungsgründe und -folgen nicht versichert zu haben.

---

38 Das voluntative Moment von Entscheidungen, d.h. das Nicht-Abwägen von Alternativen, wird zum einen mit dem Namen ‚Dezision' belegt, zum anderen in dem Modell der „resolute choice" (*McClennen* 1990), wenn auch in anderer Absicht, thematisiert: „A resolute person makes his plans on the assumption that he will follow them through" (*Sugden* 1991, 780).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Moral, können wir nun sagen, kommt nicht deshalb in die Ökonomie, weil die Ökonomie ‚höheren Normen' unterworfen werden soll, sondern weil ihr implizites moral basement in einem entscheidenden Punkt brüchig wird: und zwar dort, wo die kognitive Entscheidung der rationalen Wahlhandlung nicht garantiert werden kann. Weil die rational choice-Theorie nicht ausschließen kann, daß die ihr zugrundeliegenden Motive emotionaler Natur sein können, muß sie es hinnehmen, daß sie ihre normative Unterstellung aufgeben muß, die Individuen könnten für alle ihre Handlungen verantwortlich gemacht werden.

Die ökonomische Theorie der Rationalität enthält keine Theorie darüber, in welchem Maße ein Individuum überhaupt die Grenzen seiner Verantwortlichkeit erkennen kann. In diesem Sinne muß es — normativ — *alles* verantworten können, was es entscheidet. Das Kriterium der impliziten Moral ist aber nicht als zusätzliches Entscheidungskriterium offengelegt. Es gibt keine Instanz innerhalb der rational choice, in der das Individuum darüber entscheidet, ob es die Folgen seines Handelns verantworten kann. Indem es handelt, handelt es de lege rational.

Die moralische Tonart dieses Satzes soll nur die moralische Implikation der rational choice besonders hervorheben, zugleich die starke Anforderung an die Individuen. Hier setzt — wie schon mehrfach betont — die Moral ein. Und zwar nicht als (implizite — und zugleich stoische —) Vernunftmoral, die den Individuen Mißerfolge auszuhalten lehrt, sondern als gewöhnliche Moral, die sich der Leiden der Mißerfolger annimmt.

Moral beobachtet die Individuen nicht dort, wo sie Erfolg haben (und sich ihre ‚Rationalität' bestätigen), sondern dort, wo sie Mißerfolg haben an ihrer Rationalität. Sie gibt ihnen eine Sprache, die günstigenfalls gelingen läßt, sich selber als Person weiterhin beschreiben zu können, wo sie als Individuum rational versagt haben. *Es geht in der Moral schlicht darum, die ‚Anerkennung als Person' herzustellen oder beizubehalten*[39], *ohne diesen Vorgang vom Erfolg (oder Mißerfolg) der rational actions abhängig zu machen.*

Dieses Argument beruht auf folgender Unterscheidung: Nicht die Handlung, die jemand ausführt, ist sinnvollerweise ‚moralisch' zu nennen, sondern die

---

[39] Möglicherweise hat die rational choice-Konzeption Konstruktionsmängel — jedenfalls dort, wo es um das Verhältnis von Moral und Emotion geht: Sie kann Emotionen, Affekte etc. nur rational verarbeiten. ‚Emotional choices' können per definitionem nicht gelten. Doch gelten sie empirisch, vor allem in moralischen Zuständen, die uns hier besonders interessieren: z. B. Selbstwertgefühle, Gefühle der Achtung, der Scham, der Empörung, Lob und Tadel etc. (vgl. dazu *Tugendhat* 1993, S. 33 f.). Das Gefühl der Achtung für sich selbst z. B. ist eine moralische Konstituante, die rational kein Individuum entscheiden kann, da 1. niemand selbst über die Achtung, die andere ihm gewähren, entscheiden kann, und da 2. die Entscheidung über Achtungsgefühle etc. eine Entscheidung über den Akteurszustand ist, der die Entscheidung, die er ausführt, beeinflußt. Wir haben es mit einem Selbstreferenz-Phänomen zu tun, das üblicherweise in rational choice-Theorien nicht enthalten ist.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Person, die sie ausführt. Der Gebrauch des Wortes ‚moralisch' ist selbst dann, wenn er prädikativ erfolgt, implizit attributiv (*Tugendhat* 1993, S. 35). Die Aussage, daß jemand in einer bestimmten Handlung moralisch gehandelt hätte, meint prinzipiell, daß die Person moralisch zu nennen ist, die diese Handlung ausführt (und deshalb ausführt, weil sie eine ‚gute' oder ‚schlechte' Person ist).

Diese Unterscheidung zwischen Handlung / Person ist hilfreich, um das Verhältnis von Moral und Ökonomie zu klären. Wenn wir von ‚rationalen Handlungen' reden, ist das Individuum zwar Träger der Handlung, ansonsten aber irrelevant, da die rational choice-Theorie Selbstreferenz nicht thematisiert. Anerkennungsvorgänge können nur über die Bewertung von (erfolgreichen) rationalen Handlungen erfolgen. Wenn wir aber — mit E. Tugendhat — die Person-Identität als konstitutives Moment der Moral nennen, wäre die Moral schlecht beraten, ihre identitätsstiftende Rolle abhängig zu machen von den kontingenten Handlungserfolgen rational agierender ökonomischer Akteure.

Unsere Argumentation ging in die entgegengesetzte Richtung. Es liegt nahe, daß die Moral die Identität bildet, nicht das rationale Handeln. Da die rational choice-Theorie den ‚Identitätsverlust' der Individuen bei Enttäuschung aus Mißerfolgen (rationalen Handelns) nicht thematisieren kann, findet die Moral ihre genuine Aufgabe innerhalb der Ökonomie darin, die Handlungsfähigkeit der Individuen zu erhalten: das heißt nichts anderes, als die Identität der Person unabhängig von den Ergebnissen der rational actions zu erhalten. Das leistet sie nicht durch Informationsoptimierung, d.h. durch Verbesserung der Entscheidungskompetenz, sondern durch eine andere Thematisierung des *Selbstwertgefühls,* indem sie in einer moralischen Sprache andere Bedeutungen zur Geltung kommen läßt, *in denen die Individuen sich selbst anzuerkennen lernen können* (gegen ihre Interpretation in terms of rational choice, versagt zu haben).

In dem Maße, wie die Theorie der rationalen Wahlhandlung das Individuum als Träger gelingender Handlungen konstruiert, verschließt sie sich systematisch dem ‚Leid', das aus den Mißerfolgen kommt. Um im Mißerfolg die Handlungsfähigkeit aufrechtzuerhalten, bedarf es nicht allein normativer Unterstellungen der Identität des Individuum als jederzeit wieder handlungsfähigem Akteur, sondern von zusätzlichen identitätserhaltenden Operationen, die die Ökonomie nicht zu ihrem Objektbereich zählt, wohl aber — traditionellerweise — die Moral.

Sie erst sichert die — ansonsten nur normativ zugeschriebene — ‚Identität' der Individuen. Die Bezeichnung ‚Identität' ist hier philosophisch unaufwendig gebraucht. Sie ist nur ein anderer Name für die Erhaltung der Handlungsfähigkeit der Akteure über die Zeit. Moral sichert die Individuen gegen die aus Enttäuschung über Mißerfolge bewirkte Handlungsunfähigkeit. Ob sie dann — über die Anerkennung als Person in ihrer Handlungsfähigkeit erhalten — rational handeln oder regelhaft (oder nach anderen Kriterien), bleibt für die Klärung des Verhältnisses von Rationalität, Moral und Emotion sekundär. Folglich ist die

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Moral, so sie sich um die Personidentität in schwierigen Situationen kümmert, eine conditio sine qua non der Funktionsfähigkeit von Ökonomie.

Man könnte sagen, daß die rational choice-Theorie erst dann dem Problem der Moral adäquat begegnet, wenn sie ihren Akteuren erlauben könnte, die zu tätigenden rational choices danach auszusuchen, ob und inwieweit sie die Person-Identität mit herzustellen in der Lage sind. Dann würde die implizite Moral der rational choice, die normativ setzt, daß die Individuen verantwortlich für sich selbst seien, nicht als ‚vernunftphilosophisches a priori' eingeführt, sondern als ein Ermessensvorgang, der die rational choices danach aussucht, ob sie dem Verantwortungspostulat genügen können.

Die Pointe dieser Überlegungen besteht darin, die eigene Handlungsfähigkeit (und deren Aufrechterhaltung) zum entscheidenden Kriterium jeder Wahlhandlung zu machen. Der Akteur ist aufgefordert, in jeder Entscheidung mit zu überlegen, wie sein Akteurs- bzw. sein Handlungskompetenzstatus tangiert wird. In der älteren, utilitaristischen Handlungstheorie wurde diese Selbstreferenz thematisiert, indem sie die Akteure implizit aufforderte, nur solche Entscheidungen zu vollziehen, die ihr ‚Glück' vermehrten.

Ich glaube weder, daß diese Konzeption leichthin pratikabel ist, noch daß wir utilitaristische Lösungen unbedarft übernehmen könnten. Es kommt vornehmlich darauf an, das Person-Bewußtsein bzw. das Selbstwert*gefühl* unabhängig von den Handlungsresultaten zu halten. Da das ‚Individuum' der rational choice nur als Ordnungsmetapher von Emotionen konstruiert ist, können Emotionen (Enttäuschungen), die aus dem rationalen Ordnungsvorgang selbst entstehen, nur wieder als zu ordnende thematisiert werden. Wenn das mißlingt, fallen die betreffenden Personen als ‚irrational' aus dem ökonomischen Handlungsgefüge heraus. Die optimistische Implikation der rational choice Theorie, daß jedes ‚Individuum' jederzeit wieder rational handeln könne, läßt sich nicht über normative Zuschreibungen, sondern nur über moralische Anerkenungsprocedere herstellen.

Der Sinn dieser Darstellung beruht darauf, daß nur Individuen, die mit sich selbst identisch sind, so urteilsfähig sind, wie es die theory of rational choice voraussetzt. Folglich ist diese Identität herzustellen, nicht nur zu postulieren. Die normative Theorie der Rationalität kann ohne explizite moralische Konfundation nicht länger aufrechterhalten werden (da die implizite, gleichsam ‚natürliche' Unterstellung menschlicher Vernunft inklusive der moralischen Selbst-Verantwortlichkeit auf kulturellen Voraussetzungen des 19. Jahrhunderts beruht, die wir heute, an der Schwelle zum 21. Jahrhundert , nicht mehr zu teilen vermögen; nach Freud — so läßt sich sagen — kann die ‚Identität' von Handlungssubjekten nicht mehr ‚linear' gedacht werden. Auch Handlungstheorien müssen kontingenzfreundlicher werden, wie es im ‚multiple-self'-Ansatz begonnen wurde [*Elster* 1988]).

Das hat die Konsequenz, daß sich die normative Theorie der Rationalität weiterhin vertreten läßt, wenn sichergestellt ist, daß es hinreichend moralische

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Instanzen gibt, die die Personidentität herstellen oder erhalten können. Man kann dieses Resultat normativ wenden und für Moral plädieren, um die Funktionsfähigkeit einer auf rational actions beruhenden Ökonomie zu sichern.

Man kann allerdings auch andere Schlüsse ziehen: Wenn auf der einen Seite die moderne Ökonomie, so sie sich auf die rational choice stützt, ohne moral basement unbegründet ist, und wenn auf der anderen Seite die moral basements kleiner oder brüchig werden, dann bleibt nicht etwa die Rationalität übrig, sondern sie wird im selben Maße kleiner oder brüchig. Wir müßten uns nach einer anderen Handlungstheorie umsehen; die jetzige enthält zuviel kantianische Dichotomie von ‚Vernunft' und „Gefühl".

Literatur

Anand, P.: Philosophy of Intransitive Preference, in: The Economic Journal, Vol. 103, 1993, S. 337 ff.

Berger, L. A.: Self-interpretation, attention, and language: implications for economics of Charles Taylor's hermeneutics, in: Lavoie, D. (1991) (ed.): Economics and Hermeneutics, London / New York, 1991, S. 262 ff.

Binmore, K.: Modeling rational Players, Part I, in: Economics and Philosophy, Vol. 3, 1987.

Birnbacher, Dieter: Der Utilitarismus und die Ökonomie, in: Biervert, B. / Held, K. / Wieland, J. (Hrsg.): Sozialphilosophische Grundlagen ökonomischen Handelns, Ffm., 1990, S. 65 ff.

Ciompi, L.: Affects as central Organizing and Integrating Factors. A new Psychological / Biological Model of the Psyche, New York 1993 (a).

— Die Hypothese der Affektlogik, in: Spektrum der Wissenschaft, Februar 1993 (b), S. 76 ff.

Coleman, J. S.: Norm-Generationg Structures, in: Cook / Levi 1990, S. 250 ff.

Cook, K. S. / *Levi,* M. (eds.): The Limits of Rationality, Chicago / London, 1990.

Crouch, R.: Human Behavior: An Economic Approach, North Scituate, Mass. 1979.

Davidson, D.: Psychologie als Philosophie, in: derselbe, Handlung und Ereignis, Ffm. 1990, S. 321 ff.

— Paradoxes of Irrationality, in: Moser 1990, S. 449 ff.

Ellsberg, D.: Risk, ambiguity, and the Savage axioms, in: Moser 1990, S. 87 ff.

Elster, J.: Sour Grapes, Cambridge 1983.

— (ed.): The Multiple Self, Cambridge 1988.

— When Rationality Fails, in: Cook / Levi 1990, S. 19 ff.

Esser, H.: Alltagshandeln und Verstehen, Tübingen 1991.

Fishburn, P. C.: Nonlinear Preference and Utility Theory, Brighton 1988.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Frank, Robert: Passions within Reason, New York / London 1988 (deutsch: Strategie der Emotionen, München 1992).

Frey, B. S. / *Eichenberger,* R.: Zur Bedeutung entscheidungstheoretischer Anomalien für die Ökonomik, in: Jahrbücher für Nationalökonomie und Statistik, Bd. 206, 1989, S. 81 ff.

Fusfeld, D. R.: Rationality and Economic Behavior, in: Methodus, Vol. 4, Nr. 2, 1992, S. 63 ff.

Gähde, U.: Zur Anwendung wissenschafstheoretischer Methoden auf ethische Theorien: Überlegungen am Beispiel des klassischen Utilitarismus (in diesem Band).

Gauthier, D.: No Need for Morality: The Case of the Competetive Market, in: Philosophic exchange, Vol. 3, Nr. 3, 1982, S. 41 ff.

Hardin, R.: The Social Evolution of Cooperation, in: Cook / Levi 1990, S. 358 ff.

Harsanyi, J.C.: Morality and The Theory of Rational Behaviour, in: Sen / Williams 1982, S. 39 ff.

Hauschildt, J.: Entscheidungsziele: Zielbildung und Problemlösung, in: Witte, E. / Hauschildt, J. / Grün, O. (Hrsg.): Innovative Entscheidungsprozesse, Tübingen 1988, S. 59 ff.

Heap, S. H.: Rationality in Economics, Oxford 1989.

Heiner, R. A.: The origin of Predictable Behavior, in: American Economic Review, Vol. 73, 1983, S. 560 ff.

— The necessitiy of Imperfect Decisions, in: Journal of Economic Behavior and Organization, Vol. 10, 1988, S. 29 ff.

Hutter, M. / *Teubner,* G.: Der Gesellschaft fette Beute: Homo juridicus and homo oeconomicus als kommunikationserhaltende Fiktionen, in: Sociologica Internationalis, Sonderheft 1994.

Kahnemann, D. / *Slovic,* P. / *Tversky,* A.: Judgement under uncertainty: Heuristics and biases, Cambridge 1982.

Kersting, W.: Moralphilosophie und Dezisionismus, in: Ethik und Sozialwissenschaften, Vol. 3, H. 1, 1992, S. 23 ff.

Kliemt, H.: Papers on Buchanan and related subjects, München 1990.

Kliemt, H. / *Zimmerling,* R.: Quo vadis Homo oeconomicus? Teil 1, in: Homo oeconomicus, Bd. X (1), S. 1 ff.; Teil 2, in: Bd. X (2) 1993, S. 167 ff.

Klopstech, A. / *Selten,* R.: Formale Konzepte eingeschränkt rationalen Verhaltens, in: Todt, H. (Hrsg.): Normengeleitetes Verhalten in den Sozialwissenschaften, Berlin 1984, S. 11 ff.

Lachmann, L. M.: Austrian Economics: a hermeneutic approach, in: Lavoie, D. (ed.): Economics and hermeneutics, London, 1991, S. 134 ff.

Lavoie, D. (ed.): Economics and Hermeneutics, London / New York 1991.

McClennen. E.: Rationality and Dynamic Choice, Cambridge 1990.

Moser, P. K. (ed.): Rationality in Action, Cambridge 1990.

Nicolaides, P.: The Impossibility of Consistent Preferences: Methodological Implications, in: Methodus, Vol. 2, Nr. 2, 1990, S. 56 ff.

Priddat, B. P.: Ökonomische Knappheit und moralischer Überschuß, Hamburg 1994.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Rector, R. A.: The economics of rationality and the rationality of economics, in: Lavoie 1991, S. 195 ff.

Rorty, R.: Kontingenz, Ironie und Solidarität, Ffm. 1989.

Schnabl, H.: Bildung und Diffusion von Erwartungen, Manuskript, Institut für Sozialforschung, Abteilung Mikroökonomie, Universität Stuttgart 1987.

Schütz, A.: Das Wählen zwischen Handlungsentwürfen, in: derselbe, Gesammelte Aufsätze, Band 1, Den Haag 1971, S. 77 ff.

Selten, R.: Bounded rationality, in: JITE / Zeitschrift für die gesamte Staatswissenschaft, Vol. 146, 1990, S. 649 ff.

Sen, A. / *Williams,* B. (eds.): Utilitarianism and beyond, Cambridge 1982.

Shoemaker, S.: Rationalität und Selbstbewußtsein, in: Kienzle, B. / Pape, H. (Hrsg.): Dimension des Selbst, Frankfurt / M. 1991, S. 19 ff.

Simon, H. A.: A Behaviorial Theory of Rational Choice, in: Quarterly Journal of Economics, Vol. 69, S. 99 ff.

— Models of Bounded Rationality, Cambridge (Mass.) 1982.

— Reason in Human Affairs, Stanford 1983.

Slote, M.: Rational dilemmas and rational superogation, in: Moser 1990, S. 465 ff.

Sugden, R.: Rational Choice: A Survey of Contributions from Economics and Philosophy, in: The Economic Journal, Vol. 101, 1991, S. 751 ff.

Taylor, M.: Cooperation and Rationality: Notes on the Collective Action Problem and its Solution, in: Cook / Levi 1990, S. 222 ff.

Tugendhat, E.: Die Rolle der Identität in der Konstitution der Moral, in: Edelstein, W. / Nunner-Winkler, G. / Noam, G. (Hrsg.): Moral und Person, Frankfurt / M. 1993.

Tversky, A. / *Kahneman,* D.: Rational Choice and the Framing of Decisions, in: Moser 1990, S. 40 ff.

Vilks, A.: Neoklassik, Gleichgewicht und Realität. Eine Untersuchung über die Grundlagen der Wirtschaftstheorie, Heidelberg 1991.

Wieland, J.: Formen der Institutionalisierung von Moral in der Unternehmung (in diesem Band).

Witt, U.: Wirtschaftlicher Wandel, Entscheidungsfreiheit und die Rolle des politischen Ökonomen, in: Möschel, W. / Streit, M. / Witt, U. (Hrsg.) (1994): Marktwirtschaft und Rechtsordnung, Baden-Baden 1994, S. 225 ff.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Humboldts Sozialstaats-Skepsis

Eine spieltheoretische Rekonstruktion

Von *Rainer Hegselmann,* Bremen*

„Social theory . . . is a guide to social construction, no less than Newtonian mechanics is a guide to architecture and civil engineering. Social theory, with its more complex and inherently reactive subject matter, faces a much more difficult task."

James S. Coleman, 1992, 118

„Die Hebel und Schrauben der Lebensordnung sind gar sonderlicher und feiner Art."

Otto Neurath, 1919, 236

Im Jahre 1792 verfaßte Wilhelm von Humboldt seine Schrift *Ideen zu einem Versuch, die Gränzen der Wirksamkeit des Staates zu bestimmen.* Die Arbeit argumentiert für ein Staatsverständnis, das man heute *liberal,* vielleicht auch *libertär* nennen würde. Vor dem Hintergrund der weitgehenden staatlichen Regulierungen seiner Zeit muß man Humboldts Schrift als ein *radikales Deregulierungsmanifest* betrachten.[1] Die verschiedenen Felder möglicher Staatstätigkeit analysierend, plädiert Humboldt für einen *minimalen Staat.* Er schreibt:

> „Der Staat enthalte sich aller Sorgfalt für den positiven Wohlstand der Bürger, und gehe keinen Schritt weiter, als zu ihrer Sicherstellung gegen sich selbst, und gegen auswärtige Feinde notwendig ist; zu keinem anderen Endzwecke beschränke er ihre Freiheit" (*Humboldt* 1792, 90).

Zur Begründung seiner Ablehnung *wohlfahrtsstaatlicher* Einrichtungen führt er aus:

> „Wie jeder sich selbst auf die sorgende Hülfe des Staates verlässt, so und noch weit mehr übergiebt er ihr das Schiksal seines Mitbürgers. Dies aber schwächt die Teilnahme, und macht zu gegenseitiger Hülfsleistung träger" (a.a.O. 75).

---

* Auf die grundlegende Idee dieser Arbeit bin ich durch Gespräche mit Hartmut Kliemt gestoßen. Einen wichtigen Hinweis verdanke ich Dieter Birnbacher. Bei beiden bedanke ich mich.

[1] Humboldt veröffentlichte zunächst lediglich einige Kapitel der Schrift — vermutlich aus Gründen politischer Vorsicht. Es ist aufschlußreich, daß der 200. Jahrestag der Abfassung dieses libertären Manifests in der Bundesrepublik praktisch nicht beachtet wurde.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Humboldt hat damit lange vor Entstehen der modernen Wohlfahrtsstaaten eine These vorgelegt, die zum argumentativen Kern heutiger Wohlfahrtsstaat-Skepsis gehört. In moderner Terminologie könnte man die These Humboldts etwa wie folgt formulieren:

- Eine sozialstaatliche Daseinsvorsorge zerstört die Netzwerke solidarischer Selbsthilfe. *[Humboldtsche These]*[2]

Im folgenden werde ich diese These *in einem spieltheoretischen Rahmen, also innerhalb des Rational-Choice-Paradigmas* untersuchen. Dazu wird zunächst ein Spiel eingeführt, das elementare Formen von Solidarbeziehungen zu analysieren erlaubt (1.). Dieses Spiel soll *Solidaritätsspiel* heißen. Im Anschluß werden sozialpolitische Eingriffe so konzeptualisiert, daß deren Effekte spieltheoretisch analysierbar sind (2.). Zugleich wird eine Engführung der folgenden Analysen motiviert. Abschnitt 3 stellt die im folgenden wichtigen Eigenschaften *iterierter* Solidaritätsspiele zusammen. Im Anschluß werden die Konsequenzen sozialpolitischer Maßnahmen für *iterierte* Solidaritätsspiele untersucht (4.). Abschnitt 5 zieht Konsequenzen für die Humboldtsche These.

1. Das einfache Solidaritätsspiel

In einem *ersten* Analyse-Schritt werde ich Solidarbeziehungen durch ein *Solidaritätsspiel* modellieren. Das Spiel ist ein *2-Personen-Spiel. Netzwerke* von Solidarbeziehungen kann man aus solchen 2-Personen-Spielen *zusammensetzen.*[3] Das Spiel wird in *extensiver Form* durch [1] gegeben. Beide Spieler werden mit der gleichen Wahrscheinlichkeit p hilfsbedürftig.[4] Entsprechend ist $(1-p)$ die Wahrscheinlichkeit dafür, *nicht* hilfsbedürftig zu werden. Den ersten Zug des Spiels macht Spieler 0, nämlich die *Natur*. Gemäß der Wahrscheinlichkeit p entscheidet der Zufall darüber, ob keiner, beide, nur Spieler 1 oder nur Spieler 2 hilfsbedürftig werden. Es kann also sein, daß niemand hilfsbedürftig ist oder auch beide zugleich hilfsbedürftig werden. Im ersten Fall braucht niemand Hilfe,

---

[2] Eine prominente Rolle spielt das Humboldtsche Argument in der *ordo-liberalen* und *marktradikalen Wohlfahrtsstaat-Kritik.* Eine viel beachtete und dieser Linie folgende Arbeit ist die von Charles Murray verfaßte sozialpolitische Studie „Losing Ground — American Social Policy 1950-1980" (*Murray* 1984). Man findet die Thesee auch im Zusammenhang einer *anarchistischen* Staatskritik. Vgl. *Taylor* 1987.

[3] Vgl. *Hegselmann* 1994 und 1994 a. Dort werden die Spieler in einer 2-dimensionalen, zellulären Welt angeordnet. Jeder Spieler kann Nachbarn im Norden, Süden, Westen und Osten haben, mit denen er simultan und unabhängig voneinander Solidaritätsspiele spielt. Natürlich kann man nicht jede Art von Solidarnetzwerk auf diese Weise aus Dyaden zusammensetzen.

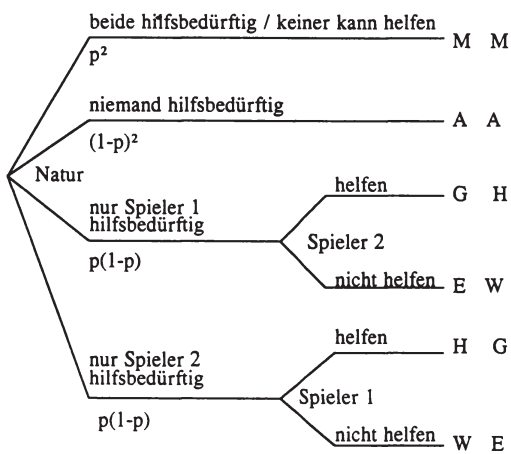
[4] Eine *gleiche* Wahrscheinlichkeit ist hier deshalb angenommen, um die Analyse *so einfach wie möglich* zu halten. Ungleiche Wahrscheinlichkeiten bieten sich an, um *Ungleichheiten von Individuen* zu modellieren. Eine Analyse der Effekte, die dann eintreten, wenn Individuen ungleich sind *und* sich vorteilsorientiert Solidarpartner suchen, wird in *Hegselmann* (1994, 1994a) gegeben.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

im zweiten Fall kann niemand helfen. Wer *selbst* hilfsbedürftig ist, kann dem anderen prinzipiell *nicht* helfen. Wer hingegen selbst *nicht* hilfsbedürftig ist, muß sich entscheiden, ob er einem hilfsbedürftigen Mitspieler hilft oder nicht. Je nachdem, wie der Zug des Spielers 0 ausgeht, ist daher im Anschluß entweder keiner, Spieler 1 oder aber Spieler 2 am Zuge.

Hinsichtlich der Auszahlungen sei angenommen, daß Hilfe zu erhalten für den Bedürftigen mit dem Nutzen *G* („*G*erettet werden") verbunden sei. Im Hinblick auf die späteren Analysen ist es sinnvoll, die Auszahlungen Hilfsbedürftiger, die keine Hilfe erhalten, obwohl andere Mitspieler, da selber nicht hilfsbedürftig, helfen könnten, *zu unterscheiden* von dem Payoff hilfsbedürftiger Spieler, die von ihren Mitspielern keine Hilfe erwarten können, da diese selber hilfsbedürftig sind. Im ersten Fall sei der Payoff *E* („*E*rtrinken")[5] , im zweiten Fall sei er *M* („*M*ist"). Ebenfalls mit Blick auf spätere Analysen wird unterschieden *zwischen* den Auszahlungen bei einer nicht gegebenen Hilfe, die auch nicht benötigt wurde, *und* der Auszahlung im Falle einer unterlassenen Hilfeleistung gegenüber einem hilfsbedürftigen Mitspieler. Im ersten Fall sei die Auszahlung *A* („*A*lles o.k."), im zweiten Fall sei sie *W* („*W*eitergehen"). Wer helfen kann und dies auch tut, der erhalte die Auszahlung *H* („*H*elfen"). Naheliegenderweise gelte als *konstitutive Bedingung* eines Solidaritätsspiels, daß $G > E$ und $W > H$.

[1] *SOLIDARITÄTSSPIEL*



BEDINGUNGEN:
$G > E$ und $W > H$
$0 < p < 1$

[5] Die aus mnemotechnischen Gründen gewählten Variablenbezeichungen sind in einer Hinsicht irreführend: Wer ertrinkt, für den ist das Spiel endgültig vorbei, wem hingegen in einem Solidaritätsspiel nicht geholfen wird, für den gibt es in dem später betrachteten Superspiel durchaus noch weitere Spielperioden.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Für das folgende ist auch eine Charakterisierung des Spiels in *strategischer Form* hilfreich.

[2]

	Solidarisch [Kooperativ]	Unsolidarisch [Defektiv]
Solidarisch [Kooperativ]	R R	S T
Unsolidarisch [Defektiv]	T S	P P

Zeilen- und Spaltenspieler stehen dabei vor der gleichen Strategienalternative. *Solidarisch* heiße eine Strategie, die dem hilfsbedürftigen anderen zu helfen vorsieht, falls man in die Lage kommt, dies zu können. Diejenige Strategie, die für eine solche Situation vorsieht, nicht zu helfen, heiße *unsolidarisch*. Die in eckigen Klammern stehenden Strategien *Kooperativ* und *Defektiv* sind deshalb aufgeführt, weil — wie sich weiter unter zeigen wird — das Solidaritätsspiel unter bestimmten Bedingungen ein Gefangenen-Dilemma mit eben dieser Strategienalternative ist.[6]

Für den Erwartungsnutzen, der den Auszahlungen T, R, P, S des Solidaritätsspiels entspricht, gilt auf Basis von [1] offenbar:

[3]
$$T = (1-p)^2 A + p^2 M + p(1-p) W + p(1-p) G$$
$$R = (1-p)^2 A + p^2 M + p(1-p) H + p(1-p) G$$
$$P = (1-p)^2 A + p^2 M + p(1-p) W + p(1-p) E$$
$$S = (1-p)^2 A + p^2 M + p(1-p) H + p(1-p) E$$

Mit Blick auf [3] sieht man sofort, daß wegen der für ein Solidaritätsspiel konstitutiven Bedingung $W > H$ immer gilt, daß $T > R$ und $P > S$. An Hand der durch [2] gegebenen strategischen Charakterisierung des Solidaritätsspiels kann man sich daher leicht klarmachen, daß die Strategie, unsolidarisch zu sein, einzige beste Antwort auf beliebige Strategiewahlen des jeweils anderen ist. Nicht solidarisch zu sein, ist also eine *dominante* Strategie. Zu einem *Gefangenen-Dilemma* wird das Solidaritätsspiel genau dann, wenn die für ein *PD* konstitutive Bedingung

[4]
$$T > R > P > S$$

erfüllt ist. Unter Rückgriff auf [3] und [4] läßt sich leicht beweisen:

---

[6] Die Payoff-Variablen sind die zumeist bei der Darstellung von *PD*-Situationen verwandten. Sie werden aus mnemotechnischen Gründen gewählt: *T*emptation, *R*eward, *P*unishment, *S*ucker's Payoff.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Hilfssatz 1:

Ein Solidaritätsspiel ist ein *PD* gdw. $(G - E) > (W - H)$. *[PD-Bedingung]*

Nennt man $(G - E)$ den *Solidaritätsgewinn* und $(W - H)$ die *Solidaritätskosten,* dann gilt also: *Ein Solidaritätsspiel ist ein Gefangenen-Dilemma genau dann, wenn für beide Spieler der Solidaritätsgewinn die Solidaritätskosten übersteigt.* Oder anders ausgedrückt: *Ein Solidaritätsspiel ist ein Gefangenen-Dilemma genau dann, wenn wechselseitige Solidarität für beide Spieler lohnenswert wäre.*

In einem Solidaritätsspiel ist es also *immer* eine *dominante* Strategie, unsolidarisch zu spielen. Zu einem Gefangenen-Dilemma wird das Spiel genau dann, wenn die Wahl der dominierten Strategien beide Spieler besser gestellt hätte. Ein Gefangenen-Dilemma ist dabei *das* klassische Beispiel einer *sozialen Falle:* Beidseitige Unkooperativität ist *Resultat der Wahl dominanter Strategien, Maximinpunkt und einziges Gleichgewicht.* Beidseitige Defektion ist damit einzige plausible Lösung des Spiels. Genau diese Lösung ist jedoch *suboptimal bzw. ineffizient,* denn es existiert eine Alternative, die beide Spieler besser gestellt hätte, nämlich beidseitige Kooperation bzw. wechselseitige Solidarität. Eben diese Lösung ist aber *für rationale Spieler nicht erreichbar,* da sie einer in die Situation eingebauten Verführung erliegen. Diese Verführung besteht darin, daß jeder weiß, daß beidseitige Solidarität zwar eine gute, eigene Unsolidarität bei gleichzeitiger Solidarität des anderen jedoch eine noch bessere Lösung wäre.

2. Eine abstrakte Modellierung sozialpolitischer Eingriffe

Für die Analyse sozialpolitischer Interventionen werde ich eine *extrem abstrakte* Sicht staatlicher Interventionen einführen: *Sozialpolitische Maßnahmen sollen aufgefaßt werden als wie auch immer herbeigeführte Veränderungen spielkonstitutiver Parameter, nämlich der Auszahlungen des Solidaritätsspiels.* Positive oder negative Steuern, Abgaben oder monetäre Transfers, Strafen irgendwelcher Art, Prämien, Bonus- oder Malus-Regelungen usw. — all dies wird auf einen einzigen Aspekt reduziert, nämlich den, die auf dem Niveau kardinaler Nutzenfunktionen gegebenen Auszahlungen des Solidaritätsspiels zu verändern. Die Frage danach, welche Effekte politische Maßnahmen für Solidarnetzwerke haben, wird so zu der Frage nach den *Konsequenzen politisch induzierter Payoff-Modifikationen für die Eigenschaften und insbesondere die Lösungen der daraus resultierenden Spiele.* Der damit aufgeworfenen Frage möchte ich unter drei Einschränkungen nachgehen:

(a) Im Prinzip kann man für jede der 6 verschiedenen Auszahlungen des Solidaritätsspiels untersuchen, welche Folgen sozialpolitisch induzierte Änderungen der Payoffstruktur haben. Ich möchte diese Untersuchung hier allerdings *beschränken* auf die Effekte von Änderungen der Auszahlungen für *Hilfsbedürftige, die keine Hilfe erhalten. Inhaltlich wird die leitende Fragestellung also*

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

zugespitzt auf das Problem einer möglichen Beeinträchtigung von Solidarbeziehungen durch eine staatliche Sozialpolitik, die im Bedürftigkeitsfalle eingreift. Dabei sollen drei verschiedene Sozialpolitiken analysiert werden, nämlich eine Politik, die die Situation derjenigen verändert, die keine Hilfe erhalten,

- weil die Mitspieler, da selber hilfsbedürfig, *nicht helfen können;*
- weil die Mitspieler, obwohl selber nicht hilfsbedürftig, *nicht helfen wollen;*
- weil die Mitspieler nicht helfen wollen *oder* nicht helfen können.

Für die weitere Analyse lege ich daher Funktionen T^*, R^*, P^* und S^* zugrunde, die jeweils aus T, R, P und S aus [3] dadurch hervorgehen, daß M um Beträge ΔM ([5], 1. Spalte) und E um Beträge ΔE ([5], 2. Spalte) verändert werden. Mit T, R, P und S als den Statusquo-Auszahlungen (also den Auszahlungen für ΔM, $\Delta E = 0$) ist durch die ersten beiden Spalten von [5] jeweils eine Menge von vier Auszahlungsfunktionen in Abhängigkeit von Änderungen um einen Nutzenbetrag Δ an den jeweils betrachteten Auszahlungen M und E gegeben. Die Funktionen ergeben sich dabei in trivialer Weise aus [3]. In der der dritten Spalte von [5] sind diejenigen Auszahlungen angegeben, die sich dann ergeben, wenn man *zugleich M und E* um Beträge ΔM und ΔE, mit $\Delta M = \Delta E = \Delta U$ verändert.

[5]

	ΔM	ΔE	ΔU
$T^*(\Delta)$	$T + p^2\Delta M$	T	$T + p^2\Delta U$
$R^*(\Delta)$	$R + p^2\Delta M$	R	$R + p^2\Delta U$
$P^*(\Delta)$	$P + p^2\Delta M$	$P + p(1-p)\Delta E$	$P + p\Delta U$
$S^*(\Delta)$	$S + p^2\Delta M$	$S + p(1-p)\Delta E$	$S + p\Delta U$

(b) Es sollen lediglich *solche* Solidaritätsspiele untersucht werden, die eine *PD-Struktur* haben. Vor dem Hintergrund von Hilfssatz 1 bedeutet dies: Die Frage nach einer möglichen sozialpolitischen Zerstörung von Solidarnetzwerken wird auf solche Solidarbeziehungen eingeschränkt, die *für alle Beteiligten vorteilhaft* sind. Die *obere* Grenze des offenen Intervalls, innerhalb dessen die sozialpolitischen Payoff-Modifikationen an der *PD*-Struktur des Solidaritätsspiels *nichts* ändern, ist gegeben durch:

[6] $$\Delta E, \Delta U < (G - E) - (W - H).$$ [7]

[7] Um diese Grenze zu ermitteln, ist lediglich zu berechnen, wann die Funktionen $T^*(\Delta)$, $R^*(\Delta)$, $P^*(\Delta)$ und $S^*(\Delta)$ aus [5] jeweils die Bedingung $T^*(\Delta E) > R^*(\Delta E) > P^*(\Delta E) > S^*(\Delta E)$ bzw. $T^*(\Delta U) > R^*(\Delta U) > P^*(\Delta U) > S^*(\Delta U)$ erfüllen. Hingegen können beliebige ΔM-Beträge eine gegebene *PD*-Struktur nicht verändern.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Eine *untere* Grenze für Payoff-Modifikationen um Beträge ΔE bzw. ΔU, bei deren Unterschreiten die *PD*-Struktur verloren ginge oder aber die konstitutiven Bedingungen $G > E$ und $W > H$ verletzt würden, gibt es *nicht.*

(c) Die Beantwortung der leitende Frage ist für ein *einfaches, d.h. nicht iteriertes* Solidaritätsspiel mit oder ohne *PD*-Struktur *nicht wirklich interessant:* An welchen Payoffs ein Sozialplaner drehen mag: *Wenn* ein Spiel ein *einfaches, nicht-iteriertes Solidaritätsspiel im Sinne von* [1] *ist,* dann ist einzig mögliche Lösung beidseitig *unsolidarisches* Verhalten. Das dabei realisierte Nutzenniveau steigt mit positiven und sinkt mit negativen ΔE-, ΔM- bzw. ΔU-Beträgen. Für die Humboldtsche These bedeutet dies: Beziehungen wechselseitiger Hilfe können im Hinblick auf *einfache* Solidaritätsspiele durch staatliche Maßnahmen *weder* geschaffen *noch* zerstört werden, denn sie sind *überhaupt nicht möglich.* Man kann ggf. allerdings dafür sorgen, daß *nach* geeigneten Änderungen der Auszahlungsstruktur das Defektionsgleichgewicht *effizient* wird. Letzteres kann dann geschehen, wenn das resultierende Spiel ein Solidaritätsspiel ohne *PD*-Charakter ist. Es gibt auch Änderungen der Auszahlungsstruktur, die wechselseitige Solidarität zu einem Gleichgewicht machen. Das derart transformierte Spiel würde dann aber *kein Solidaritätsspiel im definierten Sinne mehr sein.*

Im folgenden werde ich daher der Frage nachgehen, welche Folgen es hat, wenn in *iterierten* Solidaritätsspielen mit *PD*-Charakter Auszahlungen von Spielern, die aus dem einen oder anderen Grund keine Hilfe erhalten, durch sozialpolitische Eingriffe verändert werden. Interessant ist diese Frage deshalb, weil seit langem bekannt ist (vgl. *Taylor* 1976 und 1987; *Axelrod* 1984; *Friedman* 1986 und 1991), daß iterierte *PD*-Spiele (sog. *Superspiele*) *unter bestimmten Bedingungen* Gleichgewichte aus Superspielstrategien haben können, die vom Effekt her auf durchgängige Kooperation in allen Basisspielen hinauslaufen und in diesem Sinne *kooperativ auflösbar* sind — ein Resultat, das sich natürlich im Prinzip auch auf iterierte *Solidaritätsspiele mit PD-Charakter* übertragen lassen muß.

3. Iterierte Solidaritätsspiele

Kooperative Superspielgleichgewichte existieren dann, wenn das Spiel *unendlich* oft wiederholt wird und die Payoffs zukünftiger Spielperioden nicht zu stark abdiskontiert werden. Sei α mit $0 < \alpha < 1$ der sog. *Diskontparameter,* dann wird der Payoff einer Spielperiode t mit dem Wert α^t abdiskontiert. Bei einer derartigen exponentiellen Abdiskontierung zukünftiger Auszahlungen konvergiert der Superspielpayoff, verstanden als die Summe der (abdiskontierten) Payoffs aus einzelnen Spielperioden, gegen einen endlichen Wert. Kooperative Gleichgewichte existieren dann, wenn der Diskontparameter einen bestimmten Schwellenwert erreicht. Die Annahme, daß das Spiel unendlich oft wiederholt wird, mag man mit Blick auf die reale Welt als niemals erfüllbar ansehen. Als Ausweg bietet sich an, den Diskontparameter als *Wahrscheinlichkeit* eines nächsten Spiels zu

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

interpretieren und von Superspielen auszugehen, deren *Ende ungewiß* ist. In dieser — *von mir bevorzugten* — Deutung, die *formal* mit dem ersten Ansatz allerdings *äquivalent* ist, gewichtet man die Payoffs zukünftiger Perioden mit den Wahrscheinlichkeiten, sie überhaupt zu erreichen, und identifiziert den Superspielpayoff mit der Summe der gewichteten Payoffs der Basisspiele. Unter dieser Interpretation existieren kooperative Superspielgleichgewichte unter der Voraussetzung, *daß ein nächstes Spiel jeweils hinreichend wahrscheinlich ist.* (Es sei betont, daß eine *beliebig hohe,* im vorhinein bekannte, *endliche* Anzahl von Spielwiederholungen nach dem bekannten Backward-Induction-Argument *keine* kooperative Lösung möglich macht.)

Fragt man sich, welche *Superspielstrategien* eine kooperative Lösung möglich machen, so ist klar, daß eine Strategie, die bedingungslos und immer kooperiert (ALL C), *nicht* als Gleichgewichtsstrategie in Frage kommen kann, da bedingungslose Defektion (ALL D) immer eine bessere Antwort auf ALL C wäre als diese Strategie selbst. Umgekehrt ist beidseitige Wahl von ALL D zwar immer ein Gleichgewicht, aber eben kein kooperatives. Eine entscheidende Voraussetzung für die Existenz kooperativer Lösungen ist, daß Superspielstrategien *bedingt* sind, d.h. ihr eigenes Verhalten vom Verhalten des anderen abhängig machen können und daher z.B. zu Vergeltung fähig sind. Zwei gut untersuchte Strategien, die diese Eigenschaft haben, sind die TIT-FOR-TAT- und die TRIGGER-Strategie (vgl. *Friedman* 1986, *Taylor* 1976). TIT-FOR-TAT spielt im ersten Spiel kooperativ und macht im Anschluß immer das, was der Gegenspieler im letzten Zug tat. So wird Kooperation mit Kooperation und Defektion mit Defektion beantwortet. TIT-FOR-TAT (Abk.: TFT) ist eine sehr versöhnliche Strategie, denn auch nach einer ganzen Reihe von Defektionen des anderen ist ein TFT-Spieler schon nach einer einzigen Kooperation des anderen zu erneuter Kooperation bereit. Demgegenüber ist TRIGGER (Abk.: TR) eine völlig unversöhnliche Strategie: Schon die erste Defektion des anderen bestraft sie mit ewiger Defektion. TR beginnt allerdings das Superspiel sehr freundlich, nämlich kooperativ, und bleibt dies auch bis zur ersten Defektion des anderen.

Angewandt auf das iterierte Solidaritätsspiel möchte ich unter der *TR-Strategie* diejenige Superspielstrategie verstehen, die dem anderen so lange und falls möglich im Bedürftigkeitsfalle hilft, wie der andere nicht in einer Situation, in der Hilfe benötigt worden wäre und er sie hätte geben können, seine Hilfeleistung verweigert. Nach einer solchen unterlassenen Hilfeleistung ist ein TR-Spieler niemals wieder zu einer Hilfeleistung bereit. — Ein *TFT-Spieler* verhält sich bis zur ersten Defektion des anderen ebenso wie der TR-Spieler. Im Gegensatz zu diesem wird ein TFT-Spieler jedoch nach einer oder auch mehreren Defektionen des anderen erneut zu Hilfeleistung bereit sein, falls der andere zuvor zur Kooperation zurückkehrte, also in einer Situation, in der er helfen konnte und seine Hilfe benötigt wurde, in der Tat half.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Hilfssatz 2:

In einem iterierten Solidaritätsspiel sind Paare aus TFT- bzw. TR-Strategien jeweils Gleichgewichte gdw. für beide Spieler gilt:

$$\alpha \geq 1/[1-p(1-p)+p(1-p)(G-E)/(W-H)] = \alpha_+.$$ [8]

Nennt man die Schwellenwert-Bedingung des Hilfssatzes 2 *KOOP-Bedingung*, dann kann man also insgesamt sagen: Die *PD*-Bedingung formuliert, wann wechselseitige Solidarität *beidseitig vorteilhaft* wäre. Die KOOP-Bedingung gibt hingegen eine Mindestwahrscheinlichkeit für die Stabilität einer Interaktion an, bei deren Erreichen das, was vorteilhaft wäre, auch *machbar* wird. Man beachte, daß α_+ durch die Größen *A* und *M nicht* beeinflußt wird.

Wenn im folgenden von *Solidarbeziehungen* oder auch von *solidarischen Lösungen* die Rede ist, dann sind damit immer *iterierte* Solidaritätsspiele gemeint, die unter den jeweils gegebenen Randbedingungen *hinreichend stabil* sind, um im Superspiel kooperative Gleichgewichte möglich zu machen. Ich werde davon *ausgehen, daß immer dann, wenn die Bedingungen für kooperative Gleichgewichte erfüllt sind, solidarischen Lösungen realisiert werden.* Da die Schwellenwerte für die Gleichgewichte im Falle iterierter Solidaritätsspiele[9] zusammenfallen, kann offen gelassen werden, ob die Spieler die solidarischen Lösungen auf Basis von TFT- *oder* auf Basis von TR-Strategien[10] realisieren. Sind die Bedingungen für kooperative Gleichgewichte aus TFT- bzw. TR-Strategien hingegen *nicht* erfüllt, dann soll davon ausgegangen werden, daß beide Spieler *unsolidarisch* spielen.

4. Pfadsprünge und kritische Transfers im iterierten Solidaritätsspiel mit PD-Charakter

Im Hinblick auf die Schwellenwertbedingung aus Hilfssatz 2 stellt sich nun die entscheidende Frage, wie sich politisch induzierte Payoff-Modifikationen um Beträge ΔM, ΔE und ΔU auswirken. In [7a] bis [7c] werden die entscheidenden Effekte dargestellt. Angenommen ist, daß bestimmte Werte für *p, G, E, W, H,*

---

[8] Der Beweis ist etwas kompliziert und liegt bisher lediglich als unveröffentlichtes Manuskript vor. Die Komplikationen ergeben sich letztlich daraus, daß die Spieler bei dem hier leitenden Verständnis von TR- und TFT-Strategien in Perioden, in denen beide oder aber keiner in Not geriet, auch nicht „sehen", ob der andere kooperiert oder defektiert. Kooperation und Defektion werden vielmehr erst in Situationen offenbar, in denen genau ein Spieler hilfsbedürftig ist.

[9] Dies gilt natürlich nicht für beliebige *PD*-Spiele. Ist $T-R = P-S$, dann fallen allerdings schon im iterierten Standard-*PD* die beiden Schwellenwerte für die Existenz eines Gleichgewichts aus TFT-Strategien zusammen. Dieser Schwellenwert ist dann seinerseits identisch mit dem für Gleichgewichte aus TR-Strategien.

[10] Nash-Gleichgewichte aus TR-Strategien wären dabei auch *teilspielperfekt.*

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

A und M gegeben sind.[11] Die Interpretationen von [7a] bis [7c] gelten *unabhängig* von spezifischen Werten, die für die Visualisierungen angenommen wurden.

In Abhängigkeit von den auf der *x-Achse* abgetragen ΔM, ΔE bzw. ΔU-Werten und den auf der *y-Achse* abgetragenen Stabilitätswahrscheinlichkeiten α (Diskontparameter) läßt sich gemäß Hilfssatz 2 bestimmen, ob die Schwellenwertbedingung für die Existenz solidarischer Lösungen erfüllt ist oder nicht. In der Senkrechten sind die Payoffs U abgetragen, die ein Spieler in jeder einzelnen Spielperiode erhält, also entweder eine R^*-Auszahlung im Falle der erfüllten, oder aber eine P^*-Auszahlung im Falle der nicht erfüllten Schwellenwertbedingung. Für die Berechnung der Auszahlungen sind die (grauschattierten) Formeln für P^* und R^* aus [5] heranzuziehen. Da das Basisspiel einen PD-Charakter hat, ist sicher, daß die R^*-Auszahlungen höher sind als die P^*-Auszahlungen. Der Status quo unter Payoff-Gesichtspunkten ist durch $\Delta M, \Delta E, \Delta U = 0$ *und* irgendeine Stabilitätswahrscheinlichkeit α_0 gegeben.

(a) ΔM-Politik

Die Auszahlung M erhalten die Spieler eines Solidaritätsspiels gemäß [1] genau dann, wenn beide hilfsbedürftig geworden sind und sich daher nach Voraussetzung nicht wechselseitig helfen können. Die Effekte sozialpolitischer Interventionen, die zu Änderungen an der Auszahlung M um Beträge ΔM führen, sind in [7a] dargestellt. Man sieht dort zwei Auszahlungsebenen, die parallel zueinander verlaufen und mit wachsendem ΔM nach rechts hin steigen. Auf der oberen Ebene liegen die R^*-, auf der unteren die P^*-Auszahlungen. Eine Inspektion der Schwellenwert-Bedingung aus Hilfssatz 2 zeigt, daß der Schwellenwert für die Existenz solidarischer Superspielgleichgewichte nicht von M abhängt. Durch Manipulation von M um beliebige Beträge ΔM wird daher an der Lösung des Superspiels nichts geändert. *Man kann also durch sozialpolitisch induzierte Transfers, die die Situationen derer verbessern (oder auch verschlechtern), die sich selbst wechselseitig nicht mehr helfen können, weder funktionierende Solidarnetzwerke zerstören noch nicht funktionierende funktionstüchtig machen.*

(b) ΔE-Politik

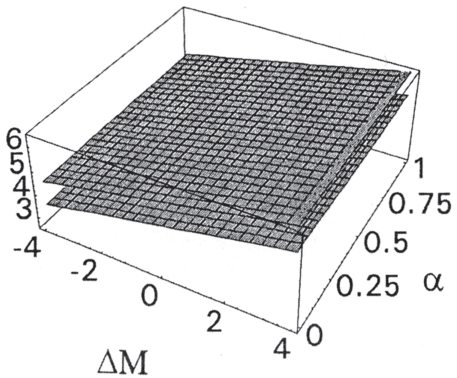
Die Auszahlung E erhält ein hilfsbedürftiger Spieler, dem von dem selber in dieser Periode *nicht* hilfsbedürftigen Mitspieler nicht geholfen wird, obwohl er dies eigentlich könnte. Durch [7b] werden die Effekte einer an der Auszahlung E ansetzenden Politik dargestellt. Ist α_0 eher niedrig, dann liegt die Status-quo-Auszahlung auf einer nach rechts hin steigenden Ebene von P^*-Auszahlungen.[12]

[11] Die 3D-Darstellungen [7a]-[7c] beruhen auf den Werten $p = 0.5$, $G = 6$, $E = 1$, $W = 8$, $H = 7$, $A = 8$ und $M = 1$.

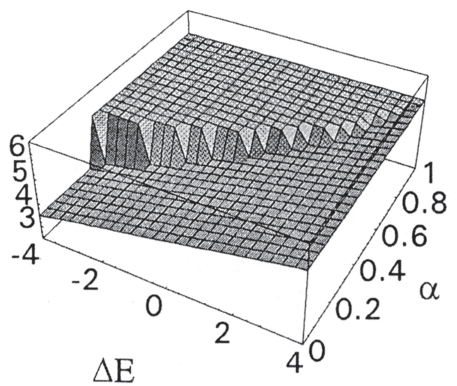
[12] Dies sollte vor dem Hintergrund von [5], 2. Spalte nicht überraschen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

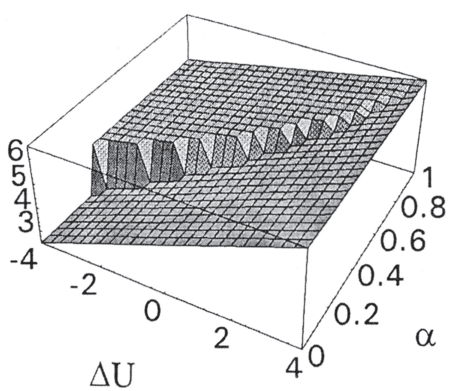
[7a]



[7b]



[7c]



DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Im Falle eines eher hohen α_0 liegt die Auszahlung auf der horizontalen Fläche von R^*-Auszahlungen. Im ersten Fall überschreitet die Stabilitätswahrscheinlichkeit nicht den erforderlichen Schwellenwert. Es ist daher nur die ineffiziente unsolidarische Lösung möglich. Im zweiten Fall wird der Schwellenwert überschritten und die mit einer höheren Auszahlung verbundene solidarische Lösung erreicht. Die obere Grenze für ΔE ist dabei in [7b] an jene Stelle gelegt, bei der das Spiel bei den gegebenen Parametern aufhören würde, ein *PD* zu sein (vgl. [6]). Ist α_0 eher niedrig, dann liegt die Status-quo-Auszahlung auf einer nach rechts hin steigenden Ebene von P^*-Auszahlungen. Im Falle eines eher hohen α_0 liegt die Auszahlung auf der horizontalen Fläche von R^*-Auszahlungen. Im ersten Fall überschreitet die Stabilitätswahrscheinlichkeit nicht den erforderlichen Schwellenwert. Es ist daher nur die ineffiziente unsolidarische Lösung möglich. Im zweiten Fall wird der Schwellenwert überschritten und die mit einer höheren Auszahlung verbundene solidarische Lösung erreicht. Die obere Grenze für ΔE ist dabei in [7b] an jene Stelle gelegt, bei der das Spiel bei den gegebenen Parametern aufhören würde, ein *PD* zu sein (vgl. [6]).

Ein etwas genauerer Blick auf [7b] offenbart *Überraschendes:* Positive Veränderungen der E-Auszahlung kann in *funktionierenden* Netzwerken in ein *Desaster* führen. Zwar schaden positive ΔE-Beträge nicht, solange sich diese Beträge *unterhalb* eines *kritischen* Wertes bewegen. *Es gibt aber offenbar positive* ΔE-Beträge, unter denen funktionierende Solidarbeziehungen zusammenbrechen, und dies, obwohl wechselseitige Solidarität nach wie vor für alle vorteilhaft wäre.[13] Hingegen können *negative* ΔE-Beträge niemals funktionstüchtige Netzwerke zerstören. Offenbar können unter dem Status quo *nicht* funktionierende Netzwerke jenseits eines kritischen *negativen* Nutzenbetrags sogar *funktionstüchtig werden.*

(c) ΔU-Politk

In [7c] finden sich die Effekte einer Politik, die *nicht differenziert* zwischen Hilfsbedürftigkeiten, bei denen wegen simultaner Bedürftigkeit niemand dem anderen helfen kann, und jenem Fall, in dem ein Mitspieler helfen könnte, es aber nicht tut. Betrachtet wird also eine Politik, die sowohl die E- wie die M-Auszahlung aus [1] um positive oder negative Beträge ΔU verändert. Auf der oberen, nach rechts hin steigenden Ebene liegen die R^*-Auszahlungen; die P^*-Auszahlungen liegen auf der unteren, aber vergleichsweise stärker steigenden Ebene.

Offenbar kann man durch positive ΔU-Beträge die Lage der Hilfsbedürftigen in *funktionierenden* Solidarbeziehungen (ähnlich wie mit einer ΔM-, jedoch im Gegensatz zu einer ΔE-Politik) zunächst verbessern. Wenn diese Beträge jedoch

[13] Denn der Zusammenbruch tritt ein, bevor das transformierte Spiel aufhört, ein *PD* zu sein. Vgl. [6].

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

einen *kritischen* Wert überschreiten, dann werden zuvor *funktionierende Solidarbeziehungen zusammenbrechen. Dies geschieht wiederum, obwohl wechselseitige Solidarität allseitig vorteilhaft wäre. Negative* Beträge senken zwar die R^*-Auszahlung, aber sie können die Funktionstüchtigkeit des Netzwerkes nicht zerstören. Wiederum können *nicht* funktionierende Netzwerke durch einen kritischen negativen Nutzenbetrag *funktionstüchtig gemacht werden.*

Eine *wesentliche Einsicht* aus [7b] und [7c] ist, daß offenbar *kritische Beträge* existieren, unter denen iterierte Spiele, die zuvor solidarisch lösbar waren, es plötzlich nicht mehr sind; umgekehrt können Spiele, die nicht solidarisch lösbar waren, unter bestimmten kritischen Beträgen solidarisch lösbar werden. Der Auszahlungpfad, auf dem die Spiele gespielt werden, springt dann vom R^*- auf den P^*- bzw. vom P^*- auf den R^*-Auszahlungspfad. Es gilt:

[8] $$\Delta E_{\text{krit}}, \Delta U_{\text{krit}} = (G-E) + (W-H)\frac{1/\alpha - 1}{p(1-p)} + 1$$

Für Sprünge vom R^*- auf den P^*-Auszahlungpfad ist der kritische Betrag dabei immer positiv, für Sprünge in die andere Richtung ist der kritische Betrag immer negativ.[14]

---

[14] Der kritische Betrag läßt sich wie folgt berechnen: Für den bei $\Delta E = 0$ zu überschreitenden Schwellenwert α_+ gilt.

[a] $$\alpha_+ = \frac{1}{1 - p(1-p)[1 - (G-E)/(W-H)]}$$

Bei einem Betrag ΔE ergibt sich der neue Schwellenwert α_+^* mit

[b] $$\alpha_+^* = \frac{1}{1 - p(1-p)[1 - (G-(E+\Delta E)/(W-H)]}$$

Das modifizierte Spiel ist solidarisch auflösbar genau dann, wenn $\alpha > \alpha_+^*$. Mit [b] ergibt sich:

[c] $$\alpha > \alpha_+^* \Leftrightarrow \Delta E < (G-E) - (W-H)\left(\frac{1/\alpha - 1}{p(1-p)} + 1\right) = \Delta E_{\text{krit}},$$

also

[d] $$\alpha > \alpha_+^* \Leftrightarrow \Delta E < \Delta E_{\text{krit}}.$$

Man sieht leicht, daß ΔE_{krit} *immer* kleiner ist als die obere Grenze des *PD*-Bereichs (vgl. [6]). Es gilt also

[e] $$\Delta E_{\text{krit}} < (G-E) - (W-H).$$

Für ΔE_{krit} kann man zeigen:

[f] $$\Delta E_{\text{krit}} > 0 \Leftrightarrow \alpha > \alpha_+.$$

Konsequenz von [d] ist, daß Spiele, die unter dem Status quo ante *solidarisch* lösbar

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Ausgelöst werden die *quasiparadoxen Effekte* dadurch, daß eine an den Auszahlungen U und E ansetzende Sozialpolitik die Solidaritätsgewinne und damit auch das Verhältnis von Solidaritätsgewinnen zu Solidaritätskosten *senkt*. Dies wiederum *erhöht den Schwellenwert* für die Stabilität der Interaktion, bei dessen Erreichen solidarische Lösungen überhaupt erst möglich sind (vgl. Hilfssatz 2). *Mit der intendierten Verbesserung des Nutzenniveaus werden also zugleich die Anforderungen an die Langfristigkeit und Stabilität von Solidarbeziehungen verschärft, und zwar möglicherweise in einem solchen Ausmaße, daß sie von einem faktisch gegebenen Solidarnetzwerk nicht mehr erfüllt werden können.* So erhalten die Individuen durch eine *nicht* nach Ursachen differenzierende ΔU-Politik in Situationen *beidseitiger* Hilfsbedürftigkeit mehr, es sind aber zugleich die Voraussetzungen dafür *zerstört*, daß im Falle *einseitiger* Hilfsbedürftigkeit der in dieser Periode nicht Hilfsbedürftige dem jeweils anderen hilft. Die Auszahlungen springen dann vom R^*- auf den niedrigeren P^*-Auszahlungspfad. *Der resultierende Erwartungsnutzen kann unterhalb des Niveaus vor jeglichem politischen Eingriff liegen.* Der gleiche zerstörerische Effekt kann auch durch eine ΔE-Politik ausgelöst werden, die ausschließlich an Situationen der Bedürftigkeit bei unterlassener Hilfeleistung ansetzt.

Sozialpolitisch induzierte Transfers *können* also offenbar in *Ineffizienzfallen* führen. Zur strategischen Reflexion fähige Mitglieder *funktionierender* Solidarnetzwerke werden daher sozialpolitische Reformen, die ihnen (und sei es sehr indirekt) vom Effekt her höhere Nutzenniveaus in Fällen garantieren, in denen sie Hilfe anderer benötigen, von den zur Hilfeleistung fähigen Mitspielern aber nicht erhalten, erst einmal sehr kritisch analysieren. *Sie würden nämlich erkennen, daß man durch solche Reformen gegenüber dem Status quo schlechter gestellt werden kann.* Und es wäre ihnen auch bewußt, daß solche Reformen vom Effekt her äquivalent einer Reform sein können, die mit *negativen Transfers* operierte, also z.B. schlicht Abgaben für diejenigen vorsähe, die Pech haben, nämlich hilfsbedürftig werden und keine Hilfe erhalten.

Umgekehrt können durch negative ΔE- bzw. ΔU-Beträge *nicht* solidarisch auflösbare Solidaritätsspiele in *solidarisch* auflösbare transformiert werden. *Unmittelbar links des Pfadsprungs ist das resultierende Nutzenniveau höher als*

---

sind, dies ab einem bestimmten ΔE_{krit} nicht mehr sind. Aus [f] ergibt sich, daß dieser kritische Nutzenbetrag *positiv* ist. Ein Spiel, das auf dem R^*-Auszahlungspfad gespielt wird, würde also durch kritische Beträge größer 0 auf den P^*-Auszahlungspfad geraten. Durch Inversion der Äquivalenz [f] ergibt sich:

[g] $$\alpha \leq \alpha_+ \Leftrightarrow \Delta E_{krit} \leq 0.$$

Spiele, die unter dem Status quo, also für $\Delta E = 0$, *nicht* solidarisch lösbar sind, würden also gemäß [g] und [d] durch ein hinreichend kleines, *negatives* ΔE solidarisch lösbar werden. — Für ΔU gilt genau die gleiche Argumentation, da der Schwellenwert durch Modifikation der M-Auszahlung nicht beeinflußt wird.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

unmittelbar rechts davon. Es kann daher sein, daß die Individuen über die Auszahlung kommen, die sie ohne jeglichen negativen Transfer erhalten hätten. Ursache dieses Effekts ist, daß negative Beträge die Relation von Solidaritätsgewinnen zu Solidaritätskosten verbessern und eben dadurch den zu erreichenden Schwellenwert für die Stabilität der Interaktion senken. Mit sinkenden Beträgen geht der Schwellenwert gegen Null. Irgendwann wird daher jede nicht funktionierende Solidarbeziehung funktionstüchtig werden. Es *kann* daher sein, daß eine Art „Steuer oder Malus auf Hilfsbedürftigkeit"[15] Mitglieder *nicht* funktionierender Solidarnetzwerke vom Effekt her gegenüber dem Status quo sogar besserstellt.

Als sozialpolitische Einsichten kann man auf Basis der Analyse von ΔM-, ΔE- und ΔU-Politiken daher jedenfalls festhalten:

— Wer Mitglieder *funktionierender* Solidarnetzwerke im Hinblick auf solche Situationen *besserstellt,* in denen sie Hilfe benötigen von anderen, die zur Hilfeleistung fähig wären, aber nicht erhalten, *riskiert den Zusammenbruch des Netzwerks* (ΔE-Politik). Das gleiche gilt für denjenigen, der Mitglieder funktionierender Solidarnetzwerke im Bedürftigkeitsfalle undifferenziert besserstellt, d.h. unabhängig davon, ob andere im Prinzip helfen könnten oder nicht (ΔU-Politik).

— Wer Mitglieder funktionierender Solidarnetzwerke *ausschließlich* in solchen Situationen besserstellt, in denen sie, da simultan hilfsbedürftig, zu wechselseitiger Hilfe nicht in der Lage sind, tut nichts, was Einfluß auf das Funktionieren von Solidarbeziehungen hätte (ΔM-Politik). Funktionierende Solidarnetzwerke können durch eine solche Politik nicht gefährdet werden.

Zugleich legen die vorangegangenen Analysen *zwei sehr allgemeine sozialtechnische Ratschläge* nahe:

— Jeder Sozialplaner sollte wissen, daß alles, was er tut, in der Regel bestimmte Anreizstrukturen schafft und verändert, auf die die Individuen dann ihrerseits reagieren. Dies wird in aller Regel heißen, daß die Individuen relativ zu den geschaffenen Payoffstrukturen rational und vorteilsorientiert ihre strategischen Möglichkeiten nutzen. Die Individuen werden dabei ein strategisches Eigenleben führen. Sozialplaner sollten daher im vorhinein zu analysieren suchen, welche Spiele mit welchen Lösungen sie vermutlich schaffen.

— Wer Anreizstrukturen gestaltet, sollte wissen, daß häufig bei Erreichen bestimmter Schwellenwerte drastische Verhaltensänderungen auftreten. Dies ist Konsequenz davon, daß sich die Gleichgewichte und Lösungen jener Spiele, die ein Sozialplaner schafft und modiziert, in der Regel *plötzlich* ändern. Politiken, die bis zu einem bestimmten Punkt erfolgreich sind, können daher jenseits bestimmter Schwellen zu einem Desaster führen. Umgekehrt können

[15] Dies könnten z.B. Kürzungen von Transferzahlungen sein.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Maßnahmen, die, bis zu einem bestimmten Punkt durchgeführt, wenig oder überhaupt nicht wirkungsvoll sind, jenseits bestimmter Schwellen sehr effektvoll sein.

Leider ist es wesentlich einfacher, diese Ratschläge zu formulieren, als sie zu beachten. So ist z.B. in der hier durchgeführten Analyse angenommen, daß die Payoffstrukturen beliebig manipulierbar sind. Eine der Hauptschwierigkeiten der faktischen Politik dürfte jedoch sein, daß sie über derartig feine Instrumente für die Gestaltung von Anreizstrukturen *nicht* verfügt. So würden sich vermutlich *monetäre* Transfers an die Spieler von Solidaritätsspielen gleich auf *alle* spielkonstitutiven Parameter auswirken. Könnte man allerdings mindestens abschätzen, in welche Richtung sie das Verhältnis von Solidaritätskosten zu Solidaritätsgewinnen insgesamt und p bzw. α verschieben, dann werden sich häufig gleichwohl die Effekte prognostizieren lassen. Ebenso kann man möglicherweise ermitteln, für welche Art von Politik man wie gute Informationen benötigt, um sie erfolgreich durchführen zu können. Im übrigen aber ist es bereits ein Gewinn zu wissen, mit welcher Art Effekten man überhaupt rechnen muß.

5. Konsequenzen für die Humboldtsche These

Nach der Humboldtschen These zerstört eine sozialstaatliche Daseinsvorsorge die Netzwerke solidarischer Selbsthilfe. Nach den hier durchgeführten Analysen ist diese These *in einer Hinsicht zu stark, in anderer zu schwach. Zu stark* ist die These deshalb, weil eine solche Zerstörung funktionierender Solidarbeziehung als Resultat sozialstaatlicher Daseinsvorsorge *zwar möglich, aber nicht notwendig* ist. Es hängt davon ab, *wie* eine Sozialpolitik im einzelnen angelegt ist. Schlecht gemacht, *kann* der von Humboldt prognostizierte Effekte allerdings in der Tat eintreten. *Zu schwach* könnte man die These deshalb nennen, weil sie außer acht läßt, daß die Individuen in der Konsequenz zerstörter Netzwerke sogar schlechter gestellt sein können, als sie es ohne solche Sozialpolitik wären. Die hier durchgeführten Analysen würden daher insgesamt eher für die folgende These sprechen:

— Es ist möglich, daß Maßnahmen im Zusammenhang einer sozialstaatlichen Daseinsvorsorge funktionierende Solidarnetzwerke zerstören. Es kann sogar sein, daß die Individuen dabei vom Effekt her schlechter gestellt werden, als sie es ohne solche Maßnahmen wären. Es läßt sich aber auch eine solche Sozialpolitik konzipieren, die die Individuen besserstellt, ohne die Solidarnetzwerke zu gefährden.

Diese modifizierte These ist ein sehr *vorläufiges* Resümee. Die Vorläufigkeit ergibt sich u.a. daraus, daß es offensichtlich zahlreiche Solidarstrukturen gibt, die durch das hier zugrundegelegte, sehr einfache Solidaritätsspiel nicht erfaßt sind. *Es wäre daher völlig verfehlt, aus den vorliegenden Ausführungen weitreichende sozialpolitische Konsequenzen zu ziehen.*

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Literatur

Axelrod, R., The Evolution of Cooperation, New York 1984; deutsche Übersetzung: Die Evolution der Kooperation, München 1987.

Coleman, J. C., The Vision of ‚Foundations of Social Theory': Analyse und Kritik 14 (1992) 117-128.

Friedman, J. W., Game Theory with Applications to Economics, Oxford 1986 [2nd Ed. 1991].

Hegselmann, R., Solidarität in einer egoistischen Welt — Eine Simulation, in: J. Nida-Rümelin (Hg), Praktische Rationalität — Grundlagenprobleme und ethische Anwendungen des Rational-Choice-Paradigmas, Berlin / New York 1994, 349-390.

— Experimental Moral Philosophy — A Computer Simulation of Classes, Cliques, and Solidarity, in: Chr. Fehige / G. Meggle / U. Wessels (Ed.), Preferences, Berlin / New York 1994 (a).

Humboldt, W. v., Ideen zu einem Versuch, die Gränzen der Wirksamkeit eines Staates zu bestimmen [1792]: Werke in fünf Bänden, Darmstadt 1980, Bd. 5, 56-233.

Murray, Ch., Losing Ground — American Social Policy 1950-1980, New York 1984.

Neurath, O., Die Utopie als gesellschaftstechnische Konstruktion [1919], in: O. Neurath, Wissenschaftliche Weltauffassung, Sozialismus und Logischer Empirismus (hgg. und eingeleitet von R. Hegselmann), Frankfurt 1979, 241-261.

Taylor, M., Anarchy and Cooperation, London 1976.

— The Possibility of Cooperation (zweite erweiterte Aufl. von Taylor 1976), Cambridge (Mass.) 1987.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Das Weniger-Ist-Mehr-Paradoxon*

Von *Wolfgang Buchholz,* Frankfurt (Oder)

1. Einleitung

Im Alltagsdenken ist die Idee, daß weniger mehr sein könne, nicht ungewohnt. In Sprichwörtern ist die Rede davon, daß Armut adele oder Mammon schnöde sei; schon in der Bibel heißt es, daß eher ein Kamel durch ein Nadelöhr gehe, als daß ein Reicher in den Himmel komme. Hinter Vorstellungen dieser Art steckt die Vermutung, daß es einen Zielkonflikt zwischen materiellem Wohlstand und den wirklich bedeutsamen Bestimmungsfaktoren des menschlichen Wohlbefindens, kurz: zwischen Haben und Sein, gibt. Es steht dabei sowohl der Genuß als auch der Erwerb von Reichtum am Pranger: Höherer Konsum gehe einher mit der Vernachlässigung nicht-materieller Glückskomponenten, wie Freude an selbstbestimmter Tätigkeit oder tieferen menschlichen Beziehungen, wie Freundschaft oder Liebe. Das Streben nach immer höherem Einkommen verstärke die Konkurrenz zwischen den Menschen und wecke Neid und Aggression. Der rücksichtslose Kampf ums Geld entzweie die Menschen. Wahre Zufriedenheit und ein harmonisches Zusammenleben ließen sich deshalb in einer bescheidenen und somit ärmeren Gesellschaft leichter erreichen.

Eine derartige Kulturkritik der Wohlstandsgesellschaft wird von Ökonomen in der Regel nicht akzeptiert. Die zentralen Gegenargumente finden sich bereits bei A. Smith. Vertraut ist v.a. seine im „Wohlstand der Nationen" artikulierte These, daß das zunächst rein egoistische Streben nach Mehrung des eigenen Einkommens gesamtwirtschaftlich positive Effekte hat. Zu den Einkommenszuwächsen, die sich bei einem aus dem individuellen Erwerbsstreben resultierenden Wirtschaftswachstum auch für die Armen ergeben, tritt eine bei steigendem Einkommen erhöhte Bereitschaft der Reichen zur Umverteilung. Wer wenig hat, kann auch wenig geben. Auch Altruismus weist die typischen Merkmale eines non-inferioren Gutes auf. Weniger bekannt hingegen sind seine Überlegungen in der „Theorie der ethischen Gefühle", in denen er das Streben nach Reichtum nicht aus dem Bemühen um Steigerung des Genusses als relativ niedrigem Beweggrund, sondern aus der verständlichen Suche nach mitmenschlicher Aner-

* Den Teilnehmern an der Sitzung des Arbeitskreises sowie W. Althammer und H. Ribhegge danke ich für wertvolle Verbesserungsvorschläge, K. A. Konrad für intensive Diskussionen über die in dieser Arbeit behandelten Themen und — last but not least — K. Tänzer für das Schreiben des Manuskripts.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

kennung zu erklären versucht: „Woher entsteht dann aber jener Wetteifer, der sich durch alle die verschiedenen Stände der Menschen hindurch zieht, und welches sind die Vorteile, die wir bei jenem großen Endziel menschlichen Lebens, das wir ‚Verbesserung unserer Verhältnisse' nennen, im Sinne haben? Daß man uns bemerkt, daß man auf uns Acht hat, daß man mit Sympathie, Wohlgefallen und Billigung von uns Kenntnis nimmt, das sind die Vorteile, die wir daraus zu gewinnen hoffen dürfen" (*Smith* 1763/1926, S. 72).

Soziale Anerkennung gilt Smith als „brennendstes Verlangen der menschlichen Natur", dessen Nichterfüllung zu tiefster Enttäuschung führt (*Smith* 1763/1926, S. 73). Daß Reichtum nicht so wichtig ist und daß man ohne die mit seinem Erwerb verbundenen Anstrengungen vielleicht ruhiger und glücklicher leben kann, stimmt dann nur bei oberflächlicher Betrachtung, nicht aber wenn man — wie A. Smith — Reichtum in der beschriebenen Weise nur als Mittel zum Zweck begreift.

Angesichts der in der Ökonomie verbreiteten positiven Einschätzung des Strebens nach Reichtum verwundert es vielleicht, daß in jüngerer Zeit in der ökonomischen Literatur vermehrt von Vorteilen, arm zu sein, die Rede ist. Insbesondere finden sich entsprechende Überlegungen in der Familienökonomie und der Theorie öffentlicher Güter. Im ersten Teil dieser Arbeit sollen zunächst die Grundgedanken, die hinter diesen Ansätzen stehen, in möglichst einfacher Form dargestellt werden. Ein Ergebnis wird dabei sein, daß in diesen Modellen *weniger* nur in einem sehr beschränkten Sinne tatsächlich *mehr* ist. Im zweiten Schritt erörtern wir deshalb die eigentlich erstaunlicheren Situationen, in denen eine vergrößerte Anfangsausstattung bei allen Beteiligten zu einem im Endeffekt für *alle* Individuen schlechteren Allokationsergebnis führt. Erst dann kann man von einem wirklichen Weniger-Ist-Mehr-(WIM)-Paradoxon sprechen.[1] Inwieweit es auf Koordinationsmängeln beruht, wird am Ende der Arbeit diskutiert.

2. Der strategische Anreiz, sich arm zu machen — eine einfache Darstellung

2.1 Das Modell der Familienökonomie

Was damit gemeint ist, daß nutzenmaximierende Individuen Anreize haben können, arm zu *bleiben* oder sich durch Verschwendung arm zu machen, soll im Rahmen eines möglichst einfachen Grundmodells herausgearbeitet werden.[2]

---

[1] Zur philosophischen Bedeutung von Paradoxien allgemein vgl. *Sainsbury* (1992), zu Paradoxien in der Ökonomie *Sinn* (1991).

[2] Vgl. im Anschluß an die rein verbale Arbeit von *Buchanan* (1975) v.a. die Modellanalysen von *Lindbeck / Weibull* (1988) und *Bernheim / Stark* (1988) sowie *Bergstrom* (1989) und *Bruce / Waldmann* (1990). Durch die folgenden Überlegungen sollen die Grundideen dieser Arbeiten aus der Familienökonomie herausgearbeitet werden. Dabei

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Wir betrachten dabei ein repräsentatives Individuum (den Empfänger), das über eine Anfangsausstattung m verfügen soll. m kann für das in Zeitstunden gemessene Arbeitspotential des Individuums oder für sein Lohneinkommen in einer ersten Periode (der „Erwerbsphase") stehen. Das Individuum hat die Möglichkeit, einen mit x bezeichneten Teil der Anfangsausstattung direkt zu konsumieren — sei es in Form von Freizeit oder aber zu Konsumzwecken in der Erwerbsphase. Der dann verbleibende Rest $m - x$ wird dazu genutzt, denjenigen Typ von Einkommen zu erzeugen, auf dessen Reduktion die strategischen Anreize abzielen. In der ersten der beiden zuvor beschriebenen Modellvarianten ist dies das Lohneinkommen, in der zweiten Modellvariante das Zinseinkommen in der jeweils zweiten Periode (der „Ruhestandsphase"). Das Niveau dieser Einkommen wird mit y bezeichnet. Bei einer konstanten Grenzrate der Transformation zwischen Anfangsausstattung und Einkommen y in Höhe von w gilt dann $y = w(m - x)$. w steht dabei entweder für den Lohnsatz oder für den um Eins vermehrten Marktzinssatz, d.h. $w = 1 + r$.

Der Nutzen des Empfängers soll von den Konsumniveaus x und y abhängen. Die Nutzenfunktion lautet also u (x, y). Es wird angenommen, daß u die üblichen Eigenschaften aufweist, d.h. in beiden Argumenten monoton wächst, zweimal partiell differenzierbar und quasi-konkav ist.

In diesem Modell bedeutet der strategische Anreiz, arm zu sein, daß das von uns betrachtete repräsentative Individuum bei seiner Nutzenmaximierung ein suboptimales Einkommensniveau y wählt. Verantwortlich hierfür sind Einkommenstransfers, die diesem Individuum von anderen Individuen gewährt werden.[3] Diese Transfers sollen vom Einkommen y abhängen, so daß wir die Transferfunktion als $T(y)$ beschreiben können. Der Einfachheit halber soll $T(y)$ differenzierbar sein. Wichtig für die folgenden Überlegungen ist $T'(y) < 0$, d.h. die Transfers sollen zunehmen, wenn das Einkommen des Empfängers sinkt. Die Transfers haben also eine kompensierende Funktion, was schon als Ad-hoc-Annahme plausibel ist.

In der Familienökonomie wird ein solcher Verlauf von $T(y)$ aus dem Nutzenmaximierungskalkül eines (oder mehrerer) altruistischer Geber *abgeleitet*. Im einfachsten Fall wird unterstellt, daß ein Geber über eine Anfangsausstattung Y verfügt, die er entweder für eigene Konsumzwecke oder aber für Transferzahlungen T an den Empfänger verwenden kann. Wenn der Geber altruistisch ist, d.h. das vom Empfänger erreichte Nutzenniveau ein Bestimmungsfaktor des eigenen Nutzens ist, lautet die Nutzenfunktion des Gebers $v(Y - T, u(x, w(m - x) + T))$

---

finden die verschiedenen Alternativen zur zeitlichen Ausgestaltung der Spielstruktur besondere Beachtung.

[3] Was speziell hinter dem Zwei-Perioden-Modell steckt, wird von *Lindbeck / Weibull* (1988, S. 1165 f.) treffend formuliert: „The source of the inefficiency is the recipient's strategic incentive to ‚squander' in an early period in order to subsequently receive more resources from the other agent, that is to ‚free-ride' on his concern.".

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

bzw. $v\left(Y-T, u(m-\frac{y}{w}, y+T)\right)$. Bei gegebenem y bestimmt sich der Wert $T(y)$ als dasjenige T, durch das dieser Nutzen maximiert wird.

Der Altruismus des Gebers zeigt sich in diesem Modell darin, daß die zweite partielle Ableitung v_2 von v positiv ist. Wenn das Einkommen y des Empfängers unter jenem Niveau liegt, das er bei isolierter Entscheidung wählen würde, sinkt sein Nutzen mit abnehmendem y. Damit vermindert sich auch der Nutzen des Gebers. Ist v_2 hinreichend groß, kann der Geber *seinen* Nutzen durch Erhöhung der Transferzahlungen steigern. $T(y)$ ist dann in der Tat eine in y fallende Funktion.

Um die Entscheidungssituation des Empfängers zu beschreiben, müssen wir $T(y)$ nicht unbedingt explizit bestimmen. Wir können auch direkt annehmen, daß $T(y)$ eine staatliche Transferzahlung (Sozialhilfe, Wohngeld) darstellt. Der Einfachheit halber gehen wir davon aus, daß das bei gegebenem $T(y)$ für den Empfänger optimale Einkommensniveau y_T als innere Lösung eindeutig bestimmt ist.

Für y_T muß dann die Marginalbedingung

$$\frac{u_1\left(m-\frac{y_r}{w}, y_r+T(y_r)\right)}{u_2\left(m-\frac{y_r}{w}, y_r+T(y_r)\right)} = w(1+T'(y_T))$$

gelten (u_1 und u_2 stehen dabei für die partiellen Ableitungen von u). Bei $T'(y_T) < 0$ ist in diesem Optimum die Grenzrate der Substitution zwischen x und y also kleiner als w , so daß die in Abbildung 1 dargestellte Konstellation vorliegt.

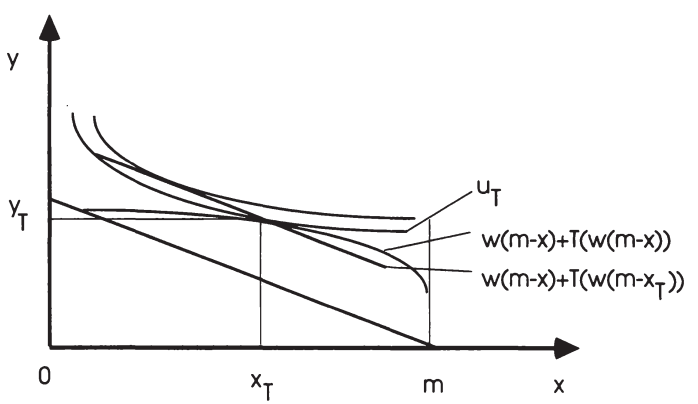


Abbildung 1

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Daß (x_T, y_r) $\left(\text{mit } x_T = m - \frac{y_T}{w}\right)$ kein *gesamtwirtschaftliches* Optimum sein kann, liegt dann auf der Hand. Dazu nehmen wir einfach an, der Transfer $T(y_T)$ würde in Form einer (vom Einkommen y unabhängigen) Transferzahlung vergeben.[4] Der Punkt (x_T, y_T) ist dann für den Transferempfänger nach wie vor erreichbar, allerdings ändert sich jetzt seine Möglichkeitskurve. Bei einer Lump-Sum-Zahlung von $T(y_T)$ entspricht die Möglichkeitskurve der Geraden durch (x_T, y_T) mit Anstieg $-w$. Bei $T(y_T) \neq 0$ erreicht der Transferempfänger dabei ein *höheres* Nutzenniveau als in (x_T, y_T). Da der absolute Betrag der Transferzahlungen aber gleich bleibt und der Nutzen des Empfängers steigt, geht hiervon ein positiver Effekt auf den Nutzen des altruistischen Gebers aus. Im Falle einer staatlichen Transferzahlung führt eine analoge Überlegung dazu, daß die Steuerbelastung und die mit ihr verbundenen Excess-Burdens nicht zunehmen.

Wie in Abbildung 1 beschrieben, kommt bei $T'(y_T) < 0$ eine Pareto-Verbesserung dann durch eine *Erhöhung* des Einkommens y zustande, d.h. das ursprüngliche Einkommensniveau y war in der Tat suboptimal.

Dabei ist angenommen, daß der Empfänger im ersten Schritt ein bestimmtes Einkommensniveau y wählt, von dem er nicht mehr abweichen kann, und der altruistische Geber im zweiten Schritt darauf reagiert. Diese Möglichkeit zur Selbstbindung bringt den Transferempfänger in die Position eines *Stackelberg-Führers,* an den sich der Geber anpaßt.[5]

Diese asymmetrische Spielstruktur ist nicht die einzig denkbare. Man könnte sich genausogut vorstellen, daß der Geber und der Nehmer simultan handeln. Das resultierende Cournot-Nash-Gleichgewicht im x-y-Diagramm des Transferempfängers muß auch jetzt auf der durch $w\,(m - x) + T(w(m - x))$ beschriebenen Kurve liegen. Allerdings ändert sich das Verhalten des Empfängers. Da dieser jetzt mit einkommensunabhängigen Transferzahlungen rechnet, ist seine Reaktionskurve im x-y-Diagramm sein Einkommensexpansionspfad, der aus allen Kombinationen von x und y besteht, in denen die Grenzrate der Substitution zwischen x und y gerade w beträgt. Das Cournot-Nash-Gleichgewicht liegt, wenn beide Güter x und y für den Empfänger non-inferior sind, bei einem im Vergleich zu y_T niedrigeren Einkommensniveau y_N und entsprechend geringeren Transferzahlungen. Der Nutzen des Empfängers ist niedriger als in (x_T, y_T). Wie sich der Nutzen des Gebers verändert, ist a priori unklar: Zwar sinkt seine Belastung durch die Transferzahlungen, andererseits vermindert über die altruistischen Präferenzen des Gebers der Nutzenverlust des Empfängers seinen Nutzen.

Man kann dieses Modell auch dadurch modifizieren, daß man für den Geber die Einkommensvariable y zu einer endogenen Größe macht.[6] Der Geber verfügt

---

[4] Vgl. zu dieser Interpretation bereits *Lindbeck / Weibull* (1988, S. 1178).

[5] Zum Zusammenhang zwischen der Möglichkeit zur Selbstbindung und der Rolle als Stackelberg-Führer vgl. *Varian* (1994).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

dann — ganz ähnlich wie der Empfänger — über eine Anfangsausstattung, die er entweder direkt konsumieren oder aber in Einkommen Y umwandeln kann, das dann entweder wieder zu Konsumzwecken oder aber zu Transferzahlungen dient. Die Nutzenfunktion des Gebers lautet dann in diesem allgemeinen Fall $v\left(M-\frac{Y}{k}, \frac{Y}{k}-T, u(m-\frac{y}{w}, y+T)\right)$, wobei der Geber seinen Nutzen durch entsprechende Wahl von Y und T zu maximieren versucht. Durch diese Verallgemeinerung ändert sich aber gegenüber dem zunächst betrachteten Modell nichts Wesentliches. Wenn der Empfänger zunächst sein Einkommen y fixiert und der Geber sich dann bei seinen beiden Entscheidungen über Y und T optimal an das vorgegebene y anpaßt, erhalten wir genauso wie zuvor eine Transferfunktion $T(y)$. Der Empfänger agiert auch in diesem Fall als Stackelberg-Führer.

Es ist jetzt aber auch eine lediglich partielle Stackelberg-Führerschaft denkbar: Für *jedes vorgegebene* Einkommensniveau Y des Gebers agiert der Empfänger zwar wie zuvor beschrieben als Stackelberg-Führer, jedoch ist es dem Geber durch die Wahl von Y möglich, sich für die darauffolgende Spielrunde zu binden. In Antizipation der zweiten Runde, in der der Spielpartner Stackelberg-Führer ist, handelt der Geber in der ersten Runde als Stackelberg-Führer. (In der dritten Spielrunde erfolgt dann der Transfer.) Man bestimmt also ein teilspielperfektes Gleichgewicht, bei dem auf beiden Stufen des Spiels die Stackelbergsche Verhaltenshypothese — allerdings mit vertauschten Rollen — zugrunde gelegt wird. Gerade im Falle von Altruismus erscheint eine derartige rationale Selbstbindung durchaus plausibel.[7] Der Geber weiß, daß ihm der andere, wenn dieser sich in die Armut hineinmanövriert hat, derart leid tun wird, daß er ihm Unterstützung auf keinen Fall verweigern wird. Der Geber weiß ferner, daß der Empfänger dies durchschaut und entsprechend strategisch handeln wird. Der Geber kann diesen von ihm selber unerwünschten Konsequenzen seiner altruistischen Präferenzen und des Handelns des Partners nur dadurch begegnen, daß er seine eigenen Handlungsmöglichkeiten begrenzt. Der Geber hält sich aus strategischer Absicht selber arm, um es sich auf diese Weise unmöglich zu machen, seinen Mitleidsgefühlen nachzugeben und einen bestimmten Betrag übersteigende Transferzahlungen zu leisten. Den Bestrebungen des Empfängers, den Geber auszubeuten, steuert dieser dadurch entgegen, daß er das Transferpotential gering hält. Man bändigt seine eigenen Emotionen (hier: das „weiche Herz") durch die Schaffung harter Fakten.

Ähnliche Überlegungen finden sich in jüngster Zeit auch in der Theorie der Sozialpolitik.[8] Wenn sich die Gesellschaft aus ethischen Gesichtspunkten ver-

---

6 Vgl. hierzu z.B. die Überlegungen von *Lindbeck / Weibull* (1988, S. 1177) und *Bruce / Waldmann* (1990, S. 160) für die jeweils von ihnen behandelten Modellvarianten.

7 Vgl. zu einer solchen rationalen Selbstbindung des Gebers insbes. *Bernheim / Stark* (1988).

8 Dieser Zusammenhang findet sich bereits bei *Lindbeck / Weibull* (1988, S. 1180 f.), wo sie darauf hinweisen, daß Zwangssparen wohlfahrtserhöhend sein kann. Vgl. auch

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

pflichtet sieht, allen ihren Mitgliedern das Existenzminimum zu sichern und entsprechende Sozialhilfezahlungen zu leisten, wird der Anreiz zur Eigenvorsorge eingeschränkt. Weshalb sollten Individuen Versicherungen für ihr Alter (Altersversicherung im eigentlichen Sinne, aber auch Krankenversicherung oder Pflegekostenversicherung) abschließen, wenn sie sich darauf verlassen können, daß die Sozialhilfe die entsprechenden Kosten trägt? Die Zahler der Sozialhilfelasten, d.h. die Mitglieder der aktiven im Erwerbsleben stehenden Generation, werden auf diese Transferverpflichtungen mit einer Einschränkung ihrer Leistung reagieren und sich auf diese Weise ärmer machen. Weder die Geber noch die Empfänger der Transfers befinden sich dann in einem von ihnen erwünschtem Zustand. Die Etablierung der gesetzlichen Alters-, Kranken- und in Zukunft eventuell einer Pflegekostenversicherung läßt sich als Versuch begreifen, im Bereich der Sozialpolitik den negativen Wohlfahrtseffekten dieses auf gesamtgesellschaftlicher Ebene auftretenden „Dilemma des Samariters" zu entgehen. Der potentielle Empfänger wird zur Eigenvorsorge verpflichtet und dadurch an dem zuvor beschriebenen strategischen Verhalten gehindert.

2.2 Das öffentliche Guts-Modell

Ein Anreiz für Individuen, arm zu bleiben, tritt aber noch in einer ganz anderen Modellwelt auf, die eigentlich viel einfacher beschaffen ist als die ursprünglich betrachtete und von vornherein auf symmetrische Cournot-Nash-Lösungen hin angelegt ist. Wir betrachten n identische Individuen, deren individuelle Einkommen y_i so gepoolt werden, daß jedes Individuum den n-ten Teil des Gesamteinkommens $\sum_{i=1}^{h} y_i = \bar{Y}$ für eigene Konsumzwecke erhält (vgl. z.B. *Holmstrom* 1982). Das Gesamteinkommen wird dann aus der Sicht der einzelnen Mitglieder des Kollektivs zu einem öffentlichen Gut. Im Cournot-Nash-Gleichgewicht, das bei wechselseitiger Anpassung der unabhängig voneinander handelnden Akteure zu erwarten ist, bleibt das Bereitstellungsniveau des öffentlichen Gutes gegenüber den optimalen Allokationen zu gering. Die Individuen erzielen also Einkommen, die unter dem für das gesamte Kollektiv optimalen Niveau liegen. Dieses Ergebnis ist schon deswegen alles andere als erstaunlich, weil es individuelle Einkommen hier letztlich überhaupt nicht gibt.

Zu einem strategischen Anreiz, arm zu bleiben, kommt es aber auch dann, wenn die Beteiligten zwar individuell zurechenbare Einkommen erzielen, diese aber dann *automatisch* gepoolt werden. Dies ist der Fall, wenn die Einkommen y_i für privaten Konsum oder aber für Beitragszahlungen zu einem öffentlichen

---

die Überlegungen von *Buchholz / Wiegard* (1992) in einem gänzlich anderen Modellrahmen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Gut verwendet werden können.[9] Die Grenzrate der Transformation zwischen privatem und öffentlichem Gut soll dabei konstant und der Einfachheit halber Eins sein.

Wir betrachten auch jetzt wieder den Fall mit n identischen Individuen mit der Nutzenfunktion

$$u(x_i, z_i, G) = u\left(m - \frac{y_i}{w}, y_i - g_i, \sum_{j=1}^{n} g_j\right).$$

g_i ist der Beitrag des i-ten Individuums zum öffentlichen Gut, $G := \sum_{j=1}^{n} g_j$ ist dessen Bereitstellungsmenge, z_i der verbleibende Konsum des privaten Gutes, nachdem der Beitrag zum öffentlichen Gut geleistet ist. Zu bestimmen ist das teilspielperfekte Gleichgewicht dieses zweistufigen Spiels. In der ersten Runde wählen die Individuen ihre Einkommensniveaus y_i, in der zweiten Runde ihre Beiträge g_i zum öffentlichen Gut. Beide Entscheidungen erfolgen völlig unkoordiniert. In der zweiten Runde ergibt sich ein Cournot-Nash-Gleichgewicht in den Beiträgen g_i bei dem jeweils gegebenen Einkommensprofil $(y_1, \ldots, y_n)$. Nach einem bekannten Resultat (*Warr* 1982, *Cornes / Sandler* 1986) ist das bei gegebenem Einkommensprofil $(y_1, \ldots, y_n)$ resultierende Cournot-Nash-Gleichgewicht bei der Bereitstellung des öffentlichen Gutes durch das aggregierte Einkommensniveau $Y := \sum_{j=1}^{n} y_j$ eindeutig bestimmt. D.h. die im Cournot-Nash-Gleichgewicht resultierenden privaten Konsumniveaus z_i^N lassen sich schreiben als $z_i^N = z_i^N(Y)$ und das Bereitstellungsniveau des öffentlichen Gutes G^N als $G^N = G^N(Y)$. Die Nutzenniveaus der Individuen, die sich bei den verschiedenen möglichen Allokationen einstellen, wenn auf der zweiten Stufe ein Cournot-Nash-Gleichgewicht resultiert, lauten

$$\bar{u}\left(m - y_i, \sum_{j=1}^{n} y_j\right) = u\left(m - y_i, z_i^N\left(\sum_{j=1}^{n} y_j\right), G^N\left(\sum_{j=1}^{n} y_j\right)\right).$$

Das aggregierte Einkommensniveau $\sum_{j=1}^{n} y_j$ ist nach Warr's Neutralitätsresultat somit selber ein öffentliches Gut, so daß sein allgemeines Resultat über die in diesem Falle resultierenden Anreize, arm zu bleiben, anwendbar ist.

Das (für alle Individuen identische) Einkommensniveau y_N, das im Cournot-Nash-Gleichgewicht resultiert, liegt unter dem optimalen Einkommensniveau y^*. Eine von y_N ausgehende simultane Einkommenserhöhung bei allen Individuen wirkt nutzenerhöhend. Dabei wird natürlich auch wieder davon ausgegangen,

---

[9] Erste Ansätze hierzu finden sich bereits in der zuvor diskutierten familienökonomischen Literatur. Vgl. *Lindbeck / Weibull* (1988, S. 1178) und *Bergstrom* (1989, S. 1149). Ein explizites Modell zu dieser Thematik findet sich aber erst bei *Konrad* (1994).

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

daß sich in der zweiten Stufe das zu den neuen Einkommenniveaus gehörige Cournot-Nash-Gleichgewicht mit dem öffentlichen Gut einstellt.

Diese Überlegungen sind nicht auf den Fall beschränkt, daß bei der Bereitstellung des öffentlichen Gutes auf der zweiten Stufe die Individuen der Cournot-Nash-Verhaltenshypothese folgen. Im Zwei-Personen-Fall etwa hängt das resultierende Gleichgewicht auch dann nur vom aggregierten Einkommen ab, wenn sich das eine Individuum in der Rolle des Stackelberg-Führers, das andere in der Rolle des Stackelberg-Folgers befindet. Nach diesem Neutralitätsresultat entspricht auch eine Entscheidung über die individuellen Einkommen einer Entscheidung über die Bereitstellung eines öffentlichen Gutes. Die entsprechende Aussage gilt, wenn bei der Bereitstellung des öffentlichen Gutes ein Gleichgewicht mit konjekturalen Variationen resultiert, sofern die diesem Gleichgewicht zugrundeliegenden Erweiterungen von einer ganz bestimmten Art sind (*Buchholz* 1993).

Den zuvor betrachteten Modellen gemeinsam ist, daß die Individuen einen Anreiz haben, ihr Markteinkommen zu beschränken. Sie haben weniger, um andere Individuen dazu zu bewegen, (ihnen) mehr zu geben. Aus diesem Blickwinkel gesehen, sind die abgeleiteten Effekte nicht mehr allzu erstaunlich. Von einem „Weniger-ist-Mehr" läßt sich im Endeffekt nicht wirklich sprechen: Den Individuen insgesamt geht es durch diese strategischen Anreize, arm zu bleiben, nicht besser, sondern schlechter. Die Fehlallokation beruht ja gerade darauf, daß insgesamt zu wenig Einkommen bereitgestellt wird. Doch auch auf individueller Ebene ist weniger nicht wirklich weniger. Diese Aussage bezieht sich ja nur auf Markteinkommen, *nicht aber auf die den Individuen ursprünglich verfügbare Anfangsausstattung,* sei es an Arbeitspotential oder an Einkommen in der ersten Periode. Wenn diese Anfangsausstattung steigt, erreichen die Individuen — auch wenn sie den zuvor betrachteten strategischen Anreizen folgen — im schließlich resultierenden Gleichgewicht ein höheres Nutzenniveau. Das gleiche gilt, wenn die Produktivität bei der Erzielung des hier interessierenden Einkommens (d.h. entweder Lohnsatz oder Zinssatz) steigt.

Im nächsten Schritt soll es darum gehen, Situationen zu beschreiben, in denen ein „Weniger-Ist-Mehr"-Phänomen in einem *strengeren* Sinne auftritt. Niedrigere Anfangsausstattungen führen dann in der Tat zu im Sinne des Pareto-Kriteriums kollektiv besseren Allokationen. Erst in einem solchen Fall wird man von einem Paradoxon sprechen können.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

3. Das strenge WIM-Paradoxon: Pareto-Verbesserungen bei reduzierter Anfangsausstattung

3.1 Ein Externalitäten-Modell

Wir betrachten zunächst ein einfaches Modell mit reziproken Externalitäten.[10] n identische Individuen mit einheitlichen Anfangsausstattungen w können diese auch jetzt wieder direkt nutzen oder aber in ein anderes Gut, dessen Menge mit x bezeichnet wird, umwandeln. Die Grenzrate der Transformation zwischen den beiden Gütern sei Eins. Der Nutzen, der für die Individuen aus dem Konsum dieser Güter resultiert, kann dann als $u\,(w - x_i, x_i)$ beschrieben werden. Zusätzlich soll der Konsum des x-Gutes aber negative Externalitäten verursachen. Diese sollen von der Gesamtmenge dieses Gutes abhängen und für alle Individuen identisch sein. Unter Einschluß dieser (negativen) Nutzenkomponente soll sich der Gesamtnutzen des i-ten Individuums dann darstellen lassen als

$$u\,(w - x_i, x_i) - \zeta\left(\sum_{j=1}^{n} x_j\right).$$

Das symmetrische Cournot-Nash-Gleichgewicht ist durch den individuellen Konsum des x-Gutes eines einzelnen Individuums vollständig beschrieben. Wenn diese Konsummenge mit x_N bezeichnet wird, muß im Cournot-Nash-Gleichgewicht die folgende Marginalbedingung erfüllt sein:

$$-u_1\,(w - x_N, x_N) + u_2\,(w - x_N, x_N) - \zeta'\,(nx_N) = 0$$

Dabei stehen u_1 und u_2 für die beiden partiellen Ableitungen von u und ζ' für die Ableitung der Schadensfunktion ζ. Die Funktionen u und ζ sollen zweifach (partiell) stetig differenzierbar sein. Der Nutzen eines Individuums im Cournot-Nash-Gleichgewicht beträgt

$$u\,(w - x_N, x_N) - \zeta\,(nx_N).$$

Wir überlegen uns, wie sich dieser Nutzen ändert, wenn die Anfangsausstattung w bei allen Individuen simultan um den gleichen Betrag dw zunimmt. Das totale Differential lautet

$$\begin{aligned} & u_1\,(dw - dx_N) + u_2 dx_N - n\zeta' dx_N \\ &= u_1 dw - (n-1)\,\zeta' dx_N - (u_1 - u_2 + \zeta')\,dx_N \\ &= u_1 dw - (n-1)\,\zeta' dx_N \end{aligned}$$

---

[10] Hiermit wird im Prinzip der Grundgedanke von *Hirsch* (1976) formalisiert. Eine explizite Fundstelle für dieses (oder ein ähnliches) Modell kann ich nicht benennen, ich vermute aber, daß es ein solches gibt.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

weil im Cournot-Nash-Gleichgewicht ja $-u_1 + u_2 - \zeta' = 0$ sein muß. (Zur Vereinfachung der Schreibweise wurden bei allen Funktionen die Argumente weggelassen.)

Eine marginale Erhöhung der Anfangsausstattung wirkt also genau dann nutzenmindernd, wenn

$$u_1 < (n-1)\,\zeta' \frac{dx_N}{dw}$$

gilt. $\frac{dx_N}{dw}$ steht dabei für die von w ausgelöste marginale Änderung der Menge des x-Gutes im Cournot-Nash-Gleichgewicht. Eine solche Konstellation ist um so wahrscheinlicher, je größer ζ' und je größer die Zahl der beteiligten Individuen ist. Beides scheint auch ohne explizite Berechnung plausibel: ζ' gibt ja den bei individuellen Entscheidungen gemäß der Cournot-Nash-Verhaltenshypothese nur partiell berücksichtigten Grenzschaden beim Konsum des x-Gutes wieder. Die Zahl der in der betrachteten Ökonomie vorhandenen Individuen ist ausschlaggebend dafür, wie stark das Cournot-Nash-Gleichgewicht von der symmetrischen optimalen Allokation abweicht, in der alle Individuen das gleiche maximale Nutzenniveau erreichen. $\frac{dx_N}{dw}$ ist bei diesen Überlegungen zunächst noch nicht näher spezifiziert. Ein expliziter Ausdruck für $\frac{dx_N}{dw}$ ergibt sich durch totale Differentiation der Marginalbedingung:

$$\frac{dx_N}{dw} = \frac{u_{11} + u_{12}}{u_{11} + u_{22} + u_{12} - n\zeta'}$$

Dabei wird $u_{12} = u_{21}$ ausgenutzt, was aus der zweifachen partiellen stetigen Differenzierbarkeit von u folgt. Wir nehmen an, daß $u_{11} < 0$, $u_{22} < 0$ sowie $u_{12} \geq 0$ gilt, woraus auch die Non-Inferiorität beider Güter für die Individuen folgt. Gilt zusätzlich $\zeta'' > 0$, ist $\frac{dx_N}{dw}$ positiv. Es ist also auch bei Berücksichtigung von Gleichgewichtseffekten möglich, daß es bei simultaner Erhöhung der Anfangsausstattungen aller Individuen zu einem Nutzenrückgang kommt.

Zur Verdeutlichung betrachten wir ein einfaches Beispiel: Wir nehmen an, daß $u\,(w-x, x) = (w-x)\,x$ und $\zeta(x) = sx$ gilt. Unter der Cournot-Nash-Verhaltenshypothese wählt dann z. B. das erste Individuum als Reaktion auf das Konsumprofil $(x_2, \ldots, x_n)$ der anderen n-1 Individuen dasjenige Konsumniveau x_1^N, bei dem

$$(w - x_1)\,x_1 - s\left(\sum_{j=1}^{n} x_j + x_1\right)$$

als Funktion von x_1 maximal wird. x_1^N ergibt sich dann als $x_1^N = \frac{w-s}{2}$.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Dieser Wert ist offensichtlich von $(x_2, \ldots, x_n)$ unabhängig, d.h. er stimmt mit dem Wert x_N im symmetrischen Cournot-Nash-Gleichgewicht überein. Die Individuen erreichen dort das Nutzenniveau

$$u^N = \left(w - \frac{w-s}{2}\right)\frac{w-s}{2} - ns\frac{w-s}{2} = (w-s)\left(\frac{w-(2n-1)s}{2}\right)$$

Nach w abgeleitet ergibt sich

$$\frac{1}{2}((w-s) + (w-(2n-1)s)) = \frac{1}{2}(2w - 2ns) = w - ns.$$

Bei $w < ns$ geht der im Cournot-Nash-Gleichgewicht erreichte Nutzen mit steigendem w zurück. Daß es allen Individuen schlechter gehen kann, wenn ihre Anfangsausstattung zunimmt, läßt sich auf die mangelnde Koordination ihres Verhaltens zurückführen. Dies wird daran deutlich, daß der im symmetrischen Optimum erreichte Nutzen mit simultan steigender Anfangsausstattung der Individuen auf alle Fälle zunehmen muß. Die Konsummenge des x-Gutes x_M in diesem Optimum bestimmt sich durch die Marginalbedingung

$$-u_1(w - x_M, x_M) + u_2(w - x_M, x_M) - n\zeta'(nx_M) = 0.$$

In das totale Differential eingesetzt, das die Veränderung des Nutzens des repräsentativen Individuums angibt, zeigt sich, daß sich bei simultaner Erhöhung der Anfangsausstattung w der individuelle Nutzen im symmetrischen Optimum gerade um $u_1 dw$ verändert, d.h. also steigt.

3.2 Ein Public-Choice-Modell

Ein Vorteil für eine Gesellschaft, arm zu sein, kann sich auch in einem Public-Choice-Modell ergeben, in dem die Entscheidung über öffentliche Projekte nach der Mehrheitsregel erfolgt. Entscheidend für das im folgenden dargestellte Argument ist, daß man es mit einer Sequenz von Projekten zu tun hat, über die nicht im Paket, sondern hintereinander abgestimmt wird.[11]

Auch jetzt soll die Ökonomie aus einer großen Zahl zunächst identischer Individuen bestehen. Diese sollen durch die folgende v. Neumann-Morgenstern-Nutzenfunktion $u(y)$ charakterisiert sein:

[11] Vgl. allgemein zu zweistufigen Mehrheitsentscheidungen *Glazer* (1989, 1993), *Glazer / Konrad* (1993) sowie *Gersbach* (1993). Die dort konstatierten Ineffizienzen werden aber nicht dazu verwendet, um kollektive Nachteile bei erhöhter Anfangsausstattung festzustellen.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

$$u(y) = \begin{cases} -\infty \text{ für } y < \underline{w} \\ y \text{ für } y \geq \underline{w} \end{cases}$$

$\underline{w}$ gibt hierbei das Existenzminimum eines Individuums an. Fällt sein Einkommen unter $\underline{w}$, verliert es seine Existenzgrundlage. Es wird deshalb niemand ein Projekt, bei dem er mit einem unter $\underline{w}$ liegenden Endeinkommen rechnen muß, akzeptieren. Gegenüber solchen Projekten verhält sich ein Individuum also extrem risikoavers.

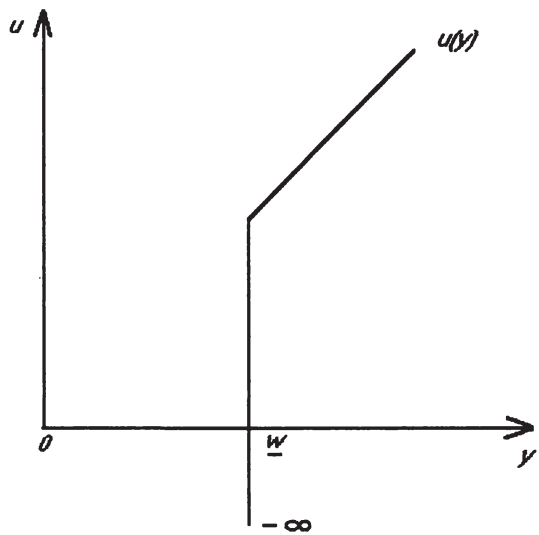


Abbildung 2

Der Einfachheit halber wird angenommen, daß sich die Individuen im Bereich $y > \underline{w}$ risikoneutral verhalten, was sich in einem linearen Verlauf der Nutzenfunktion niederschlägt.

Im ersten Fall sollen alle Individuen im Ausgangszustand über das Existenzminimumeinkommen $\underline{w}$ verfügen. Ein erstes kollektives Projekt soll den p-ten Teil der Individuen um g_1^a und den (1-p)-ten Teil der Individuen um g_1^b besserstellen.

Es sei $g_1^b > g_1^a$ und $p < {}^1/_2$, d. h. bei der insgesamt resultierenden Pareto-Verbesserung wird der höhere Einkommenszuwachs einer Mehrheit der Individuen zuteil. Ein zweites riskantes Projekt kann bei allen Individuen zu Gewinnen g_2 und individuellen Verlusten l_2 führen.

Die Verlustwahrscheinlichkkeit betrage q, die Gewinnwahrscheinlichkeit $1 - q$. Der Erwartungswert dieses zweiten Projektes sei positiv, d.h. es soll

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

$(1 - q)\, g_2 - ql_2 > 0$ gelten. Zusätzlich wird $l_2 > g_1^b$ angenommen, d.h. der beim zweiten Projekt drohende Verlust soll die im ersten Projekt erzielbaren Zusatzgewinne übersteigen. Das Ergebnis einer sequentiellen Mehrheitsentscheidung ist dann sofort klar: Das erste Projekt, bei dem sich alle besserstellen, wird (sogar einstimmig) angenommen, das zweite Projekt einstimmig abgelehnt: Alle Individuen würden bei einem für sie negativen Projektausgang unter ihr Existenzminimum fallen und dann das Nutzenniveau $-\infty$ realisieren.

Ganz anders sieht das Entscheidungsverhalten aber aus, wenn das Anfangseinkommen w_o der Individuen über $\underline{w}$ liegt. Wir nehmen an, daß

$$\underline{w} - g_1^a + l_2 > w_o > \underline{w} - g_1^b + l_2$$

gilt. Die Individuen, die sich nach dem ersten Projekt auf einem relativ höheren Einkommensniveau befinden, bleiben dann auch nach Fehlschlag des zweiten Projektes über ihrem Existenzminimum. Sie werden deshalb das zweite Projekt akzeptieren. Dieses Votum ist bei einer Entscheidung nach der Mehrheitsregel aber entscheidend, da sich aufgrund der Annahme $p < {}^1/_2$ diese Individuen in der Mehrheit befinden.

Für die Individuen, die durch das erste Projekt nur den niedrigeren Einkommenszuwachs g_1^a erzielen, ändert sich hingegen nichts. Sie sind nach wie vor davon bedroht, durch das zweite Projekt unter ihr Existenzminimum zu fallen. Nachdem das erste Projekt vorgenommen wurde, können sie das zweite Projekt aber nicht mehr verhindern, da sie dann in der Minderheit sind. In Antizipation der vorhersehbaren Entscheidung über das zweite Projekt werden die Individuen ihre Entscheidung über das erste Projekt revidieren. Um zu vermeiden, daß eine Mehrheit zugunsten des zweiten Projektes zustande kommt, durch die sie möglicherweise um ihr Existenzminimum gebracht werden, lehnen sie bereits einstimmig das erste Projekt ab. Alle Individuen bleiben dann auf dem ursprünglichen Einkommensniveau w_0. Wenn $w_0 < \underline{w} + g_1^a$ ist, stellen sich die Individuen dann im Endeffekt gegenüber der Situation, in der ihr Anfangseinkommen nur $\underline{w}$ betrug, schlechter. Daß eine Konstellation möglich ist, in der es in der beschriebenen Weise zu einem kollektiven Vorteil, arm zu sein, tatsächlich kommen kann, soll durch ein Zahlenbeispiel belegt werden.

Das erste Projekt sei durch die Gewinne $g_1^a = 2$ und g_1^b gekennzeichnet, die mit den Wahrscheinlichkeiten $p = {}^1/_4$ bzw. ${}^3/_4$ resultieren sollen. Beim zweiten Projekt sollen der Gewinn $g_2 = 8$ und der Verlust $l_2 = 4$ mit der gleichen Wahrscheinlichkeit $q = 1 - q = {}^1/_4$ auftreten. Der Erwartungswert des zweiten Projekts beträgt dann 2 und ist also insbesondere positiv.

Das Existenzminimumeinkommen $\underline{w}$ soll 10 Einheiten betragen. Wenn das Anfangseinkommen aller Individuen mit $\underline{w}$ zusammenfällt, führt das erste Projekt ¾ der Bevölkerung zum Einkommensniveau 13, ¼ zum Einkommensniveau 12. Das zweite Projekt ist dann für die Angehörigen beider Einkommensklassen

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

unattraktiv, da sie dadurch mit der Wahrscheinlichkeit ½ unter das Existenzminimum fallen. Die Individuen votieren in diesem Fall einstimmig für das erste und einstimmig gegen das zweite Projekt.

Beträgt das Anfangseinkommen jedoch $w_0 = 11{,}5$, sind die relativen Gewinner des ersten Projekts bereit, auch das zweite Projekt in Angriff zu nehmen. Ihr Einkommen nach dem ersten Projekt beträgt 14,5, so daß sie selbst bei einem Verlust von 4 mit 10,5 über dem Existenzminimum von 10 bleiben. Sie können dann die relativen Verlierer des ersten Projekts überstimmen, die nach dem zweiten Projekt mit 9,5 das Existenzminimum verfehlen würden. Um dieser Gefahr zu entgehen, verzichten die Individuen einstimmig bereits auf das erste Projekt. Die Individuen bleiben dann auf dem Einkommensniveau 11,5, d.h. sie stellen sich *alle* gegenüber dem Zustand schlechter, in dem mit 10 ihr Anfangseinkommen geringer war.

Die Ineffizienz, die auf diese Weise zustande kommt, ließe sich natürlich durch eine Umverteilungspolitik zugunsten der Verlierer des zweiten Projekts vermeiden. Wenn die Mehrheit auf jeder Stufe der alleinige Souverän ist, sind entsprechende Ankündigungen nicht glaubhaft. Es gibt ja keine übergeordnete Machtinstanz, die ein solches Versprechen gegen die Mehrheit durchsetzen könnte. Aus theoretischer Sicht möglich schiene es, daß sich die relativen Gewinner des ersten Projektes durch Bindung eines Teils ihres Zusatzeinkommens zumindest illiquide machen. Dies kann etwa durch die kollektive Finanzierung öffentlicher Güter geschehen. Im Gegensatz zur Umverteilung wirkt eine solche Festlegung glaubwürdig, weil diese Individuen aus dem öffentlichen Gut selber Nutzen ziehen. Wenn dieser Nutzen hinreichend groß ist, wirkt dieses Commitment glaubhaft.

Die vorherigen Überlegungen werfen in gewisser Weise auch ein etwas anderes Licht auf den vielfach beschworenen Zielkonflikt zwischen Effizienz und Verteilung: Die Möglichkeit größerer Abweichungen von dem ursprünglichen Gleichverteilungszustand kann ja das Zustandekommen einer effizienten Allokation verhindern, so daß hier eine Umverteilung sogar die Erreichung des Effizienzziels wahrscheinlicher macht. Beim Zusammenschluß zweier Staaten bedeutet dies nichts anderes als ein Variante der alten Behauptung, daß die *Homogenität* von Staaten für ihre Integration förderlich ist. Im Falle von nacheinander vorgenommenen Projekten wird deutlich, daß hinter der Ablehnung von Projekten, die zu großen Einkommenszuwächsen bei einem Teil der Bevölkerung führen, eventuell nicht primär Neid, sondern ein Streben nach Selbstschutz steckt. Eine egalitäre Gesellschaft liegt im Interesse von Individuen, weil ein hohes Einkommen der anderen diese zu wagemutig machen und zu Projekten animieren kann, deren Verlustrisiken gemessen am eigenen Einkommen zu hoch erscheinen.

Mit diesen Überlegungen wird auch ein weiteres Argument dafür geliefert, weshalb die ökonomische Transformation in manchen ehemals sozialistischen Ländern langsam verläuft und teilweise Rückschläge erleidet. Die Einkommens-

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

differenzen, die durch die ersten Reformschritte entstehen, schaffen unterschiedliche Interessenlagen. Die anfänglichen Nicht-Gewinner bzw. Verlierer im Reformprozeß müssen möglicherweise befürchten, durch eine von der reicheren Mehrheit beschlossene riskante Wirtschaftspolitik zusätzliche Einbußen zu erleiden, die von ihnen dann nicht mehr verkraftet werden können. Die dadurch bewirkte Zurückhaltung gegenüber einer tiefgreifenderen ökonomischen Umwandlung ist dabei um so ausgeprägter, je weniger eine Grundsicherung der Individuen garantiert scheint. Doch selbst wenn entsprechende Regelungen formell beschlossen würden, kann keiner mit Sicherheit wissen, ob sie nicht nach der nächsten Wahl wieder zurückgenommen werden. Ein Zeitinkonsistenzproblem stellt sich auch hier. Vertrauen in die Beständigkeit von Maßnahmen zur sozialen Absicherung kann nur aufgrund von Erfahrungen oder aber durch Herausbildung eines gesellschaftlichen Grundkonsensus über die Notwendigkeit von Sozialpolitik entstehen. Nach Ablauf weniger Jahre des Transfomationsprozesses kann mit beiden ernsthaft noch nicht gerechnet werden.

4. Zusammenfassung

Viele Resultate, die auf den ersten Blick paradox erscheinen, erweisen sich bei genauerer Betrachtung kaum mehr als sonderlich überraschend. Dies gilt auch für das hier betrachtete „Weniger-Ist-Mehr-Paradoxon“: Bei den ersten beiden Modelltypen hatten die Individuen deshalb einen Anreiz, arm zu bleiben, weil sie erwarten konnten, vom Einkommen anderer mitprofitieren und in diesem Sinne eine Freifahrerposition einnehmen zu können. Dies galt unabhängig davon, ob man durch direkte Transfers oder durch gemeinsame Nutzung eines öffentlichen Gutes an fremdem Einkommen teilhaben kann. Daß die Verwässerung von Eigentumsrechten (durch Einkommensumverteilung oder durch Sozialversicherungen) die Leistungsanreize vermindert, ist ein Standardargument aus der Finanzwissenschaft. Der Unterschied zu den hier betrachteten Modellen liegt primär darin, daß der Verzicht auf Einkommen durch staatlichen Zwang und nicht aufgrund von Entscheidungen gemäß der eigenen Präferenzen erfolgt.

Bei den beiden anderen diskutierten Modellen stellte sich eine Gesellschaft bei niedriger Anfangsausstattung deswegen besser, weil mit höherem Einkommen Koordinationsprobleme zunahmen oder überhaupt erst auftraten. Werden einige oder alle Individuen reicher, verursachen sie negative externe Effekte, deren Nutzeneffekt den des ursprünglichen Wohlstandszuwachses übersteigt. Korrekturmöglichkeiten für die Externalitäten waren innerhalb der Modelle nicht vorgesehen. Im eigentlichen Externalitätenmodell kann man diese Zusatzannahme dadurch rechtfertigen, daß die Einkommensvariable besser beobachtbar ist und sich damit besser kontrollieren läßt als die Konsumvariable. Im Public-Choice-Modell wurden die mit steigendem Einkommen verbundenen Nachteile durch die strikte Anwendung der Mehrheitsregel und deren gegenüber existentiellen

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Bedrohungen der jeweiligen Minderheit rücksichtslosen Anwendung begründet. Eine verbindliche Ethik, die ein solches rein egoistisches Verhalten ausschließt, könnte die resultierende Ineffizienz vermeiden helfen.

Literatur

Bergstrom, T. (1989), A Fresh Look at the Rotten Kid Theorem and other Household Mysteries, in: Journal of Political Economy 97, 1138-1159.

Bernheim, D. / *Stark* O. (1988), Altruism within the Family Reconsidered: Do Nice Guys Finish Last?, in: American Economic Review 78, 1034-1045.

Bruce, N. / *Waldman* N. (1990), The Rotten Kid Theorem Meets the Samaritan's Dilemma, in: Quarterly Journal of Economics 105, 155-165.

Buchanan, J. M. (1975), The Samaritan's Dilemma, in: E. S. Phelps (ed.), Altruism, Morality and Economic Theory, New York, 71-85.

Buchholz, W. (1993), A Further Perspective on Neutrality in a Public Goods Economy, in: Public Finance Quarterly 21, 115-118.

Buchholz, W. / *Wiegard* W. (1992), Allokative Überlegungen zur Reform der Pflegevorsorge, in: Jahrbücher für Nationalökonomie und Statistik 209, 441-457.

Cornes, R. / *Sandler* T. (1986), The Theory of Externalities, Public Goods, and Club Goods, Cambridge U.K.

Gersbach, H. (1993), Politics and the Choice of Durability: Comment, in: American Economic Review 83, 670-673.

Glazer, A. (1989), Politics and the Choice of Durability, in: American Economic Review 79, 1207-1213.

— (1993), Politics and the Choice of Durability: Reply, in: American Economic Review 83, 674-675.

Glazer, A. / *Konrad,* K. A. (1993), The Evaluation of Risky Projects by Voters, in: Journal of Public Economics 52, 377-390.

Hirsch, F. (1976), Social Limits to Growth, Cambridge (Mass.).

Holmstrom, B. (1982), Moral Hazard in Teams, in: Bell Journal of Economics 13, 324-340.

Konrad, K. A. (1994), The Strategic Advantage of Being Poor: Private and Public Provision of Public Goods, in: Economica 61, 79-92.

Lindbeck, A. / *Weibull* J. W. (1988), Altruism and Time Consistency: The Economies of Fait Accompli, in: Journal of Political Economy 96, 1165-1182.

Sainsbury, R. M. (1993), Paradoxien, Stuttgart.

Sinn, H. W. (1991), Paradoxa in der Volkswirtschaftslehre, in: Jahrbuch für Sozialwissenschaften 42, 263-277.

Smith, A. (1926), Theorie der ethischen Gefühle, herausgegeben von Walther Eckstein, Leipzig.

Varian, H. R. (1994), Sequential Contributions to Public Goods, in: Journal of Public Economics 53, 165-186.

Warr, P. (1983), The Private Provision of a Public Good is Independent of the Distribution of Income, in: Economics Letters 13, 207-211.

DOI https://doi.org/10.3790/978-3-428-48170-5 | Generated on 2025-12-18 13:04:14
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/