

ÖKONOMIE UND GESELLSCHAFT

Jahrbuch 2

*Wohlfahrt und
Gerechtigkeit*

*Campus Verlag
Frankfurt/New York
1984*

Distributive Gerechtigkeit und Wohlfahrt

**Zum Grundproblem einer utilitaristischen Ethik
und Wohlfahrtsökonomie**

Rudolf Kötter

Distributive Gerechtigkeit und Wohlfahrt gehören zu den wenigen Themen, bei denen sich im Laufe der Zeit eine intensive und fruchtbare Zusammenarbeit zwischen Philosophie und Ökonomie entwickelt hat. Das geistige Band bildet dabei der Utilitarismus, der, im 19. Jahrhundert von Philosophen entwickelt, heute hauptsächlich von Ökonomen seine systematische Weiterführung erfährt. Als leistungsfähiges Instrument für die Auseinandersetzung mit dem klassischen Utilitarismus hat sich dabei der umfangreiche formale Apparat der Ökonomie erwiesen, durch den viele Konzepte der Ethik in einem neuen und manchmal helleren Lichte erscheinen, wenngleich nicht verschwiegen werden darf, daß der Fortschritt zu größerer formaler Präzision seinen Preis hat: Unter der scharf zupackenden Begrifflichkeit der Ökonomen fiel manches aus der philosophischen Tradition als Späne zur Seite, was zumindest auf einen zweiten Blick dem Korpus „Gerechtigkeit und Wohlfahrt“ zugerechnet werden muß. Ich möchte diesem Punkt ganz besondere Aufmerksamkeit widmen, wobei es mir weniger darauf ankommt, Interpretationen zurechtzurücken, als vielmehr darauf, daß einiges, was aus dem Blickfeld geraten ist, der Diskussion wieder verfügbar gemacht wird.

Distributive Gerechtigkeit und Wohlfahrt stehen in der utilitaristischen Tradition eigentlich von Beginn an in einem eigentümlichen Spannungsverhältnis. Wenn wir, wie in der Ethik üblich, mit „gerecht im weiteren Sinne“ eine Gesellschaftsordnung bezeichnen, deren Normen durch ein Prinzip der Ethik gerechtfertigt sind, dann ist dem Utilitarismus entsprechend eine Gesellschaftsordnung gerecht im weiteren Sinn, in der eine sich aus den individuellen Nutzensvorstellungen zusammensetzende Gesamtwohlfahrt optimiert wird. Das Problem dabei ist, wie der individuelle Nutzen sich bilden soll, wie er aggregiert werden kann und wie das Optimum zu bestimmen ist. Dieser Gerechtigkeitsbegriff tritt in Konkurrenz zu einem engeren, der mit dem „Hauptsatz“ der Theorie der distributiven Gerechtigkeit verknüpft ist: Wird ein Mitglied einer Gesellschaft dann und nur dann bei einer Verteilung bevorzugt, wenn es qualifizierter ist als die anderen und erhalten gleichqualifizierte Mitglieder gleiche Teile, so ist die der Verteilung zugrunde liegende Gesellschaftsordnung gerecht im engeren Sinn. Als Grundprobleme ergeben sich hier die Fragen nach

den Kriterien, denen zufolge die Qualifikation des einzelnen bzw. die Gleichheit der Teile zu beurteilen sind.

Diese beiden Gerechtigkeitsbegriffe und die mit ihnen verbundenen Probleme stehen prima facie recht unvermittelt nebeneinander. Es hat gleichwohl in der utilitaristischen Tradition nicht an Versuchen gefehlt, Gerechtigkeit im engeren Sinne unter das Nützlichkeitsprinzip zu subsumieren. Diesen Versuchen möchte ich im folgenden nachgehen.

Die Gliederung des Aufsatzes ist einfach. Ich werde zunächst zwei Richtungen aus der Wohlfahrtsökonomie kritisch diskutieren, die ihren Ursprung unmittelbar bzw. mittelbar im Utilitarismus haben: Einmal den Neo-Utilitarismus in seiner ordinalen und kardinalen Variante und zum anderen die Theorie der Fairness, die im wesentlichen konzipiert wurde, um den Problemen der neo-utilitaristischen Ansätze aus dem Weg zu gehen. Letztere mag auch als Kontrastfolie dazu dienen, die Charakteristika utilitaristischen Denkens noch deutlicher erscheinen zu lassen.

Ausdrücklich betont sei, daß es hier nicht darum geht, einen einigermaßen umfassenden Überblick über die genannten Strömungen zu geben, vielmehr sollen lediglich die relevanten Grundzüge dargestellt werden. Und daß hierzu manchmal auch mit einem recht kräftigen Pinsel gemalt werden muß, werden hoffentlich auch die detailfreudigen Social-Choice-Theoretiker einsehen.

Schließlich möchte ich aber auch auf die philosophischen Quellen der Wohlfahrtsökonomie zurückkommen und die moderne, oft recht glatte Behandlung des Themas kontrastieren mit einigen der diffizilen und manchmal auch widersprüchlichen Ansätzen der klassischen utilitaristischen Ethik; dies in der Hoffnung, daß sich zumindest die Perspektive, unter der man unser Thema zu sehen gewohnt ist, weitet und der Sache neue Aspekte abgewonnen werden können.

1. Der Neo-Utilitarismus

1.1. Das utilitaristische Programm im allgemeinen

Die moderne Wohlfahrtsökonomie ist in ihren Hauptströmungen mehr oder weniger eng der philosophischen Tradition des Utilitarismus verbunden. Zwar baut man heute auf einer wesentlich ausgefeilteren Nutzentheorie auf als es die Klassiker taten, teilt mit ihnen jedoch die Überzeugung, daß Handlungen nach ihren Ergebnissen zu beurteilen sind und das Kriterium für ihre Beurteilung im Nutzen liegt, der aus den Folgen der Handlung für die Betroffenen ent-

springt. Insoweit ist es also gerechtfertigt, von den nutzentheoretisch fundierten Richtungen der Wohlfahrtsökonomie als von einem „Neo-Utilitarismus“ zu sprechen.

Im Rahmen des Neo-Utilitarismus selbst lassen sich nun zwei Varianten danach unterscheiden, ob die Nutzenabwägung in einem kardinalen, interpersonell vergleichbaren oder in einem ordinalen Sinne zugelassen wird. Bevor ich mich diesen Richtungen zuwende, möchte ich zunächst den allgemeinen Rahmen eines „klassischen“ utilitaristischen Programms (so wie es in der Wohlfahrtsökonomie verstanden wird) abstecken, damit deutlich wird, was das „Neue“ am Neo-Utilitarismus ist. Beginnen wir mit einigen Definitionen.

Wir betrachten eine Menge N von n Personen p_1, \dots, p_n und einen m -dimensionalen Güterraum $G \subset \mathbb{R}^m$, der zu \mathbb{R}^m strukturgleich ist. Jedes Element $g \in G$ stellt ein qualitativ und quantitativ genau spezifiziertes Güterbündel dar. Mit X bezeichnen wir die Menge der möglichen Verteilungen der Elemente $g \in G$ auf die n Personen, d.h. die Elemente $x \in X$ sind $m \times n$ -Matrizen, die als Verteilungstableaus interpretiert werden können und gewissermaßen die Folgen einer je bestimmten Verteilungspolitik für die Mitglieder aus N verzeichnen. Der Einfachheit wegen gehen wir davon aus, daß die möglichen Verteilungen auch sämtlich zulässig und realisierbar sind und z.B. keinen institutionellen Einschränkungen unterliegen.

Jedes Mitglied $p_i \in N$ möge nun eine kardinale, interpersonell vergleichbare und streng konkave Nutzenfunktion $U_i : X \rightarrow \mathbb{R}$ besitzen, die den Nutzwert angibt, den eine Verteilung $x \in X$ für p_i hat. Dies bedeutet, daß Nutzdifferenzen eines Individuums als meßbar und darüber hinaus die Nutzenfunktionen selbst als auf einen bestimmten Maßstab beziehbar angenommen werden. Um das utilitaristische Programm allgemein formulieren zu können, müssen wir zunächst ein gesellschaftliches Wohlfahrtsmaß W festlegen, das sich aus den individuellen Nutzenfunktionen zusammensetzen soll. Dazu wählen wir eine Menge S von positiven Gewichten, $S = \{ \lambda = (\lambda_1, \dots, \lambda_n) \mid \lambda_i > 0 \text{ für alle } p_i \text{ und } \sum_{i=1}^n \lambda_i = 1 \}$ und definieren für jede Verteilung $x \in X$ eine Abbildung $W(x, \lambda) : X \times S \rightarrow \mathbb{R}$ als $W(x, \lambda) = \sum_{i=1}^n \lambda_i U_i(x)$. $W(x, \lambda)$ stellt also den mit λ gewichteten Durchschnitt der individuellen Nutzenfunktionen $U_i(x)$ für eine bestimmte Verteilung dar und wird gesellschaftliche Wohlfahrtsfunktion genannt. Das „klassische“ utilitaristische Programm läßt sich jetzt kurz so formulieren (vgl. Yaari 1981, 1-39, bes. 8):

Wähle zunächst zu einer gegebenen Menge X von Verteilungen den „richtigen“ Gewichtungssatz $\lambda^* \in S$. Dann suche das $x^* \in X$, für das $W(x^*, \lambda^*) = \max !$ gilt.

Die Annahmen der Strukturgleichheit von G mit \mathbb{R}^m und der strengen Konkavität der individuellen Nutzenfunktionen sichern die prin-

zielle Lösbarkeit dieses Programms. Bei dieser Formulierung ist sofort zu sehen, daß im „klassischen“ Utilitarismus das Problem der Gerechtigkeit im engeren Sinne bei der Wahl des „richtigen“ Satzes von Gewichtungen angesiedelt ist. Die übliche Lösung dieses Problems ist denkbar einfach: Alle p_i werden lediglich bezüglich ihrer Eigenschaft, Träger einer Nutzenfunktion zu sein, betrachtet. Das heißt, die Unterschiedlichkeit der Menschen wird nur insoweit registriert, als sie sich in unterschiedlichen Nutzenfunktionen niederschlägt. Damit hat man aber keine Kriterien zur Verfügung, die es erlauben würden, eine differenzierte Struktur der λ_j zu bilden und setzt diese folgerichtig alle gleich dem Anteil des einzelnen an der Gesamtbevölkerung. Da auf diese Weise jede Nutzenfunktion mit gleichem Gewicht in die Summation eingeht, ist $W(x^*, \lambda^*) = \max!$ als Aufforderung zur Maximierung des arithmetischen Mittels der individuellen Nutzenfunktionen zu interpretieren. Es ist wenig verwunderlich, daß auf diesem allgemeinen Niveau das Thema Gerechtigkeit kaum eine ausführliche Diskussion erfahren hat, es verschwindet fast völlig hinter der schlichten und schematischen Prozedur der Durchschnittsbildung. Dennoch wird bei der formalen Darstellung deutlicher als bei der üblichen verbalen Formulierung des Nützlichkeitsprinzips, daß Gerechtigkeit im Sinne von Gleichgewichtung nicht einfach der Nutzenoptimierung inhärent ist, sondern sich offensichtlich einem eigenen Prinzip verdankt.

Bei genauerer Betrachtung des Programms fallen nun drei kritische Probleme ins Auge. Das erste hängt mit der Forderung nach kardinalen, interpersonell vergleichbaren Nutzenfunktionen zusammen. Spätestens seit Pareto hat man in der Ökonomie erkannt, daß die psychologischen Annahmen, die einst das kardinale Nutzenmaß fundieren sollten, unhaltbar sind und sich in zunehmendem Maße von dieser Forderung distanzieren. Die ordinale Nutzentheorie hat sich in der Ökonomie durchgesetzt, sie ist heute in der Mikroökonomie zur herrschenden Lehre avanciert, wenngleich gerade in der Wohlfahrtsökonomie der Gedanke an eine kardinale Nutzenfunktion lebendig geblieben ist. Es ist also zu prüfen, ob man dem utilitaristischen Programm nicht auch unter Zugrundelegung einer ordinalen Nutzentheorie gerecht werden kann, wenn auch vielleicht nur in einem abgeschwächten Sinne. Dieser Frage ist vor über 30 Jahren K.J. Arrow nachgegangen und er hat mit seiner Antwort zugleich den Grundstein für einen sich lebhaft entwickelnden Forschungszweig gelegt. Seinem Ansatz werden wir uns im nächsten Abschnitt zuwenden.

Das zweite Problem liegt in der Vermittlung zwischen den individuellen Nutzenfunktionen und der gesellschaftlichen Wohlfahrtsfunktion. Aus der Formulierung des utilitaristischen Programms ist nicht ersichtlich, wie man als Individuum oder als Träger institutionalisierter Entscheidungsbefugnis dazu kommt, seine eigene Nutzen-

funktion in eine gesellschaftliche Wohlfahrtsfunktion zu transformieren. Ist dies Ausdruck individueller Rationalität oder Folge eines moralischen Appells, und wenn letzteres, welche Bedingungen müssen gegeben sein, damit dieser Appell auf fruchtbaren Boden fallen kann?

Das dritte Problem ergibt sich schließlich aus der schon erwähnten unklaren Stellung des Gleichgewichtsgrundsatzes im Rahmen des Optimierungsprogramms. Den beiden letzten Fragen ist unter anderem J. Harsanyi nachgegangen und wir werden darauf im übernächsten Abschnitt zurückkommen.

1.2 Der Neo-Utilitarismus in seiner ordinalen Variante

Zur Darstellung der ordinalen Variante des utilitaristischen Programms müssen wir zunächst unseren formalen Rahmen etwas umändern. Wir gehen wieder aus von der Menge N der n Personen, betrachten aber jetzt nur eine endliche Teilmenge $A \subset X$ von m alternativen Verteilungen als entscheidungsrelevant. Jede Person bringt nun diese Alternativen in eine Reihenfolge, d.h. sie ordnet die Elemente der Menge A gemäß einer zweistelligen Ordnungsrelation R_i , die reflexiv, vollständig und transitiv ist. R_i wird „Präferenzordnung des Mitglieds p_i “ genannt. Man kann definieren: Für alle $p_i \in N$ und alle Paare $x, y \in A$ gilt: $x R_i y$ und $y R_i x \stackrel{\text{def}}{=} I_i y$ (Indifferenz) sowie $x R_i y$ und $\neg y R_i x \stackrel{\text{def}}{=} P_i y$ (strikte Präferenz). Die Menge aller möglichen Präferenzordnungen auf A wird mit A^m bezeichnet (es ist dies die Menge aller $m!$ Permutationen der Elemente von A). Eine Liste aller Präferenzordnungen der n Personen zu einem Zeitpunkt wird Präferenzprofil q genannt. Alle denkbaren Kombinationen von Präferenzordnungen zu solchen Listen bilden die Menge Q^n der möglichen Präferenzprofile (die Menge Q^n hat also $(m!)^n$ Elemente).

An die Stelle der kardinalen Nutzenfunktionen sind jetzt also Präferenzordnungen getreten, die sich nicht mehr addieren lassen, so daß an die Stelle von W eine andere gesellschaftliche Wohlfahrtsfunktion $R: Q^n \rightarrow A^m$ treten muß, die die Transformation der individuellen Präferenzordnungen in eine gesellschaftliche Präferenzordnung leistet. Für die Frage des Wohlfahrtsmaximums kommt dann auch nur ein ordinales Kriterium in Frage, nämlich die Anzahl derjenigen $p_i \in N$, die sich für eine bestimmte Präferenzordnung aussprechen. Wenn wir uns für einen Augenblick die bekannte – und recht unglückliche – Formulierung des utilitaristischen Prinzips: „Maximiere das Glück der größten Zahl!“ in Erinnerung rufen, dann läßt sich das „klassische“ utilitaristische Programm als Versuch werten, den Satz in Richtung „Maximiere den Nutzen einer bestimmten Anzahl!“ zu korrigieren. Im ordinalen Utilitarismus läuft der Korrekturversuch dagegen

in die andere Richtung: „Maximiere die Anzahl derer, die eine bestimmte Nutzenvorstellung (Präferenzordnung) besitzen!“.

Arrow nahm die Gedanken auf und gab vier Bedingungen für alle Kandidaten einer gesellschaftlichen Wohlfahrtsfunktion R an, welche die Ermittlung einer optimalen gesellschaftlichen Wohlfahrt im ordinalen Sinne sicherstellen sollen. Diese Bedingungen lauten²:

- 1) *Rationalität*: Die gesellschaftliche Wohlfahrtsfunktion liefert eine Ordnung, die den gleichen Bedingungen genügt wie eine individuelle Präferenzordnung, d.h. die gesellschaftliche Präferenzordnung ist reflexiv, vollständig und transitiv.
- 2) *Unabhängigkeit von irrelevanten Alternativen*: Die gesellschaftliche Ordnungsbeziehung zweier Alternativen x, y hängt einzig und allein von den individuellen Ordnungsbeziehungen dieser beiden Alternativen ab.
- 3) *Schwaches Pareto-Prinzip*: Ziehen alle Mitglieder $p_i \in N$ die Alternative x der Alternative y vor ($x P_i y$), dann muß auch die gesellschaftliche Wohlfahrtsfunktion $x P y$ ergeben.
- 4) *Diktatorausschluß*: Es gibt kein Mitglied aus N , dessen individuelle Präferenzordnung unabhängig von den Präferenzordnungen der anderen Mitglieder durch die gesellschaftliche Wohlfahrtsfunktion zur gesellschaftlichen Präferenzordnung wird.

Drei Bemerkungen mögen zur Erläuterung dieser Annahmen dienen.

I) Die Annahmen 1) und 2) bedingen, daß als Konkretisierung einer gesellschaftlichen Wohlfahrtsfunktion nur Entscheidungsverfahren in Frage kommen, bei denen die gesellschaftliche Präferenzordnung allein durch den paarweisen Vergleich der Alternativen ermittelt wird. II) Die Bedingung 2) bedeutet darüberhinaus, daß in die gesellschaftliche Präferenzordnung nicht eingehen darf, in welcher Weise zwei Alternativen relativ zu den anderen Alternativen in einer individuellen Präferenzordnung geordnet sind, ob sie also beide an der Spitze oder am Ende der Ordnung stehen oder ob zwischen sie andere Alternativen eingeordnet sind. Des weiteren wird die Möglichkeit ausgeschlossen, in die gesellschaftliche Wohlfahrtsfunktion eine unterschiedliche Gewichtung der individuellen Präferenzordnungen einzubauen: die gesellschaftliche Wohlfahrtsfunktion darf im paarweisen Vergleich nur die individuellen Ordnungen der Alternativen berücksichtigen und zwar ohne Ansehen der Person. III) Die Bedingungen 3) und 4) beinhalten den Gedanken, daß die gesellschaftliche Wohlfahrtsfunktion optimierbar sein soll in dem oben angegebenen ordinalen Sinne, jedenfalls soll nicht ein Minderheitsvotum das Votum einer Mehrheit dominieren können (Diktatorausschluß).

Es ist unschwer zu erkennen, daß das Gerechtigkeitsproblem in der ordinalen Variante des utilitaristischen Programms genauso schlicht behandelt wird, wie im obigen allgemeinen Fall. Auch hier besitzt je-

der als Träger einer Präferenzordnung gleiches Gewicht. Entscheidend ist natürlich die Frage, ob es überhaupt eine gesellschaftliche Wohlfahrtsfunktion gibt, die diesen Bedingungen genügt. Daß diese Frage nicht trivial ist, mag das folgende altbekannte Beispiel zeigen. Beginnen wir zunächst mit einer Definition. Wir wollen eine Teilmenge $D \subset N$ entscheidende Koalition nennen, wenn aus $x P_i y$ für alle $p_i \in D$ folgt $x P y$. Durch das einstimmige Votum der entscheidenden Koalition wird also die gesellschaftliche Präferenz festgelegt. Nimmt man nun das Mehrheitswahlrecht als ein Verfahren, durch das die entscheidende Koalition bestimmt wird, so kann sich bei drei Alternativen x, y, z folgendermaßen eine inkonsistente gesellschaftliche Präferenzordnung ergeben: Angenommen ein Drittel der Mitglieder aus N zieht x dem y und y dem z vor, ein weiteres Drittel y dem z und z dem x und das letzte Drittel schließlich z dem x und x dem y . Dann ergeben sich Mehrheiten für $x P y$, $y P z$ und $z P x$. Nach Transitivität gilt aber z.B. $x P z$, d.h. das Mehrheitswahlrecht erzeugt keine gesellschaftliche Wohlfahrtsfunktion, die den vier Bedingungen genügt. Arrow konnte nun dieses Ergebnis verallgemeinern. Sein Satz lautet:

Satz 1: Es gibt keine gesellschaftliche Wohlfahrtsfunktion, die den Bedingungen 1) - 4) genügt³.

Also nicht nur das einfache Mehrheitswahlsystem, sondern jedes denkbare Wahlverfahren kann zu inkonsistenten gesellschaftlichen Präferenzen führen.

Dieser Satz stellte natürlich eine ungeheuerere Herausforderung dar und bei seiner Würdigung konzentrierte sich die Kritik insbesondere auf die Annahmen 1) und 2), die nicht recht plausibel erschienen und denen man die „Hauptschuld“ an dem negativen Resultat zuschob. Warum sollte das Ergebnis eines gesellschaftlichen Entscheidungsprozesses unbedingt in der Aufstellung einer gesellschaftlichen Präferenzordnung liegen? Wie wir aus den Diskussionen um die Nutzentheorie in der neoklassischen Mikroökonomie wissen, ist die Forderung nach Vollständigkeit und Transitivität in dieser Allgemeinheit schon bei individuellen Präferenzordnungen äußerst problematisch. Und selbst wenn wir diesen kritischen Punkt ausklammern und seine Diskussion in den nächsten Abschnitt verschieben, bleibt hier festzuhalten, daß es im Grunde genommen nur zwei Fälle gibt, für die in der Erstellung einer gesellschaftlichen Präferenzordnung ein sinnvolles Entscheidungsergebnis gesehen werden kann. Einmal, wenn es um die Festlegung einer zeitlichen Reihenfolge geht, nach der bestimmte Projekte verwirklicht werden sollen; zum anderen, wenn die Verwirklichung von Alternativen von nicht beeinflussbaren Umständen abhängt. Im letzteren Fall gibt die Präferenzordnung dann an, was zu tun ist, wenn die Alternative mit der höchsten Präferenz sich nicht

verwirklichen läßt (Festlegung von zweit- und drittbesten Lösungen). Abgesehen von diesen Fällen genügt eigentlich die Festlegung einer Alternative als gesellschaftlich verbindlich und gerade unser Fall der Entscheidung zwischen alternativen Verteilungen läßt dieses Argument besonders einleuchtend erscheinen.

Diese Kritik mündet in die Forderung, anstelle der Arrow'schen gesellschaftlichen Wohlfahrtsfunktion eine gesellschaftliche Entscheidungsfunktion zu setzen, die lediglich eine Alternative als kollektiv verbindlich erklärt, wobei diese ausgewählte Alternative allerdings ebenfalls als gesellschaftlich optimal in einem utilitaristisch ordinalen Sinne interpretierbar sein muß. Formal gesehen heißt dies, daß man anstelle der Abbildung $R: Q^n \rightarrow A^m$ eine Abbildung $E: Q^n \rightarrow A$ betrachtet. Es ist nun keine Frage, daß es gesellschaftliche Entscheidungsfunktionen gibt, die diktatorfrei sind und dem schwachen Pareto-Prinzip genügen, allerdings taucht bei diesen Entscheidungsfunktionen als ein neues Problem das der Manipulierbarkeit der Präferenzen auf. Damit ist gemeint, daß jemand durch die Angabe einer anderen Präferenzordnung als der, die er eigentlich meint, sicherstellen kann, daß die gesellschaftliche Entscheidungsfunktion ein für ihn günstigeres Ergebnis liefert. Ein Beispiel mag dies verdeutlichen. Angenommen, drei Personen haben sich zwischen vier Alternativen w, x, y, z zu entscheiden. Die gesellschaftliche Entscheidungsfunktion ordnet nun zunächst dem maximalen Element in einer individuellen Präferenzordnung vier Punkte, dem zweitbesten drei Punkte usw. zu. Dann werden die Punktzahlen pro Alternative summiert und die Alternative mit der höchsten Punktzahl ausgewählt (bei Punktgleichheit entscheidet das Los). Ein mögliches Präferenzprofil ist z.B.:

P ₁	P ₂	P ₃
w	x	z
x	w	w
y	y	x
z	z	y

In diesem Falle würde die gesellschaftliche Entscheidungsfunktion die Alternative w ermitteln. Dieses Ergebnis könnte nun p_2 dadurch vereiteln, daß er anstelle seiner „wahren“ Präferenz x, w, y, z die manipulierte Präferenz x, y, z, w nennt und folglich die von ihm am meisten geschätzte Alternative x zur gesellschaftlich verbindlichen Alternative bestimmt werden würde.

Die Bedeutung dieses Problems für unsere Thematik liegt auf der Hand: Entscheidungsfunktionen, die manipulierbar sind, können nicht zur Bestimmung des gesellschaftlichen Optimums, selbst in unserem abgeschwächten Sinne, verwendet werden, da hier Minderheiten die Möglichkeit haben, durch die Wahl einer geschickten Strategie den Willen der Mehrheit zu torpedieren. Es müßten also Entschei-

dungsfunktionen gesucht werden, die strategischer sind. Gibbard und Satterthwaite haben jedoch gezeigt, daß diese Suche vergeblich ist und folgenden Satz bewiesen:

Satz 2: Es gibt keine strategischere gesellschaftliche Entscheidungsfunktion, die dem schwachen Pareto-Prinzip und der Forderung nach Diktatorausschluß genügt⁴.

Die Sätze von Arrow und Gibbard/Satterthwaite besagen nicht, daß jeder Entscheidungsprozeß in ein Dilemma geraten muß, aber die Möglichkeit des Auftretens läßt sich nicht für jeden Fall vermeiden. Allerdings scheint mit den Sätzen von Arrow und Gibbard/Satterthwaite noch nicht das letzte Wort in dieser Angelegenheit gesprochen. Ansatz für weitere Überlegungen bietet die Forderung nach Eindeutigkeit der gesellschaftlichen Entscheidungsfunktion. Gerade wenn wir Entscheidungen zwischen alternativen Verteilungen betrachten, ist es gut vorstellbar, daß unterschiedliche Verteilungen gleiche individuelle Präferenz erhalten und deshalb als Wert einer gesellschaftlichen Entscheidungsfunktion nicht nur genau ein Element aus A zugelassen werden sollte, sondern Teilmengen $E(q) \subset A$. $E(q)$ sei dabei die zu dem Präferenzprofil $q \in Q^n$ gehörige Auswahlmenge von Verteilungen, zwischen denen gesellschaftliche Indifferenz besteht. Die Betrachtung von mengenwertigen Entscheidungsfunktionen wirft nun Probleme auf, die zu einigen Re-Definitionen zwingen und die ich wenigstens ganz knapp andeuten möchte.

So nennen wir jetzt ein Mitglied $p_i \in N$ einen Diktator, wenn aus dem Umstand, daß $x P_i y$ und $x \in E(q)$ folgt, daß y nicht in die Auswahlmenge $E(q)$ gehört. Einige Schwierigkeiten macht die Definition der Manipulierbarkeit von nicht linearen Präferenzen, denn hier muß geklärt werden, was man darunter verstehen will, daß p_i eine Menge (nämlich die Auswahlmenge $E(q')$, die sich bei der manipulierten Präferenzordnung R_i' ergibt) einer anderen Menge (der Auswahlmenge $E(q)$ ohne Manipulation) vorzieht. Sind X und Y zwei Teilmengen aus A , so kann man definieren: $X P_i Y \stackrel{\text{def}}{=} \text{für alle } x \in X \text{ und } y \in Y \text{ gilt } x R_i y$ und für wenigstens ein Paar x, y gilt $x P_i y$. Es gibt hier noch eine ganze Reihe von mehr oder weniger plausiblen alternativen Definitionen, ich möchte aber auf diese Einzelheiten nicht eingehen (vgl. Gärdenfors 1979, Feldman 1979 b). Die Frage lautet also jetzt, ob es nicht-manipulierbare, diktatorfreie mengenwertige Entscheidungsfunktionen gibt. Sie läßt sich in dieser Allgemeinheit nicht beantworten, aber es lassen sich zumindest positive Teilergebnisse vermerken. Wenn wir z.B. als gesellschaftliche Entscheidungsfunktion eine Regel (die sog. „Pareto-Regel“) wählen, die den Präferenzprofilen die zugehörigen paretooptimalen Auswahlmengen zuordnen, dann läßt sich zeigen, daß diese Regel strategi-

sicher und diktatorfrei ist⁵. Es sei dahingestellt, ob die Pareto-Regel einen praktisch bedeutsamen Entscheidungsmechanismus darstellt, jedenfalls weist das Ergebnis in eine interessante Richtung.

Vor zuviel Optimismus sei jedoch gewarnt. Solange man an der schematischen Gleichbehandlung der Individuen festhält und so darauf verzichtet, dem Entscheidungsproblem eine reichere Struktur zu geben, werden Unmöglichkeitssätze das Bild der Social-Choice-Theory prägen⁶. Die Manövriermenge für sinnvolle Modifikationen der üblichen Annahmen dieser Theorie scheint mir zu gering zu sein, um auf spektakuläre Ergebnisse hoffen zu können. So hat sich also der Weg des ordinalen Utilitarismus als recht dornig erwiesen und der Satz von der Erhaltung der Schwierigkeit eines Problems wurde wieder einmal bestätigt: Resultate, die man um den Preis gravierender Annahmen erzielt hat, lassen sich nicht billiger bekommen, lediglich ein Austausch von Schwierigkeiten ist möglich.

1.3 Der Neo-Utilitarismus in seiner kardinalen Variante

Da sich in den fünfziger Jahren unter dem Eindruck des Satzes von Arrow der ordinale Ansatz in der Wohlfahrtsökonomie als wenig vielversprechend erwies, mußte man sich, wollte man nicht ganz auf eine utilitaristische Wohlfahrtsökonomie verzichten, mit erneuten Anstrengungen um eine plausiblere Begründung der kardinalen Nutzentheorie bemühen. Es ist also nur scheinbar paradox, daß gerade in einer Zeit, in der die ordinale Nutzentheorie ihre großen Triumphe feierte, das Interesse an der schon totgesagten kardinalen Nutzentheorie wieder erwachte. Angeregt wurde dieses Interesse durch den Vorschlag von J.v. Neumann und O. Morgenstern⁷, die kardinale Nutzentheorie als Mischung von ordinaler Nutzentheorie und einer Theorie rationalen Verhaltens unter Unsicherheit zu konzipieren.

Zur Skizzierung dieses Ansatzes gehen wir von der Betrachtung einer Situation aus, in der Alternativen bewertet werden müssen, deren Eintreten von bestimmten unsicheren Ereignissen abhängt. Auf unser Verteilungsproblem bezogen, bedeutet dies zunächst, daß die Verteilungen $x \in X$ als abhängig von der Wahl bestimmter Verteilungspolitiken h_1, \dots, h_k angenommen werden. Die Verteilungspolitiken bilden zusammen die Menge H , die Menge der bedingten Verteilungen $\{x_1 | h_1, \dots, x_k | h_k\}$ wird mit L bezeichnet. Alle $p_i \in N$ mögen eine Präferenzordnung R_i bezüglich der Verteilungen besitzen, die folgender Annahme genügt:

- 1) Die Präferenzordnung R_i auf X , $i \in \{1, \dots, n\}$, ist vollständig, transitiv und reflexiv, außerdem sind für jedes $\bar{x} \in X$ die Mengen $\{x \in X | x R_i \bar{x}\}$ und $\{x \in X | \bar{x} R_i x\}$ abgeschlossen.

Da unser Güterraum G eine abzählbare Basis besitzt, besitzt auch X eine solche und unter Annahme 1) ist jede Präferenzordnung R_i äquivalent mit einer ordinalen stetigen Nutzenfunktion $u_i : X \rightarrow \mathbb{R}$ (vgl. Debreu 1959, 56ff.).

Nun soll der Eintritt eines Ereignisses $x \in X$ unter einer Bedingung $h \in H$ unsicher sein in dem Sinne, daß der Entscheidungsträger $p_i \in N$ nur die Wahrscheinlichkeiten $w_j = \text{Prob}_i(h_j)$ für alle $h_j \in H$ kennt (mit Prob_i wird die Setzung einer subjektiven Wahrscheinlichkeit durch p_i bezeichnet). Als Rationalitätskriterien für die Entscheidung unter Unsicherheit werden die folgenden Postulate genommen (vgl. Harsanyi 1977a):

- 2) Sei $L = \{x_1 | h_1, \dots, x_k | h_k\}$ und $L' = \{x_1 | h'_1, \dots, x_k | h'_k\}$. Angenommen, der Entscheidungsträger $p_i \in N$ kennt die Wahrscheinlichkeiten für alle h_1, \dots, h_k ebenso wie für die h'_1, \dots, h'_k und weiß außerdem, daß für alle $j = 1, \dots, k$ gilt $\text{Prob}_i(h_j) = \text{Prob}_i(h'_j)$. Dann möge p_i indifferent zwischen L und L' sein (probabilistic equivalence postulate).

Dies bedeutet insbesondere, daß es dem Entscheidungsträger gleichgültig ist, aufgrund welcher Handlungen sich die für ihn relevanten Folgen einstellen, lediglich das Resultat zählt. Oder anders ausgedrückt: der Nutzen einer Verteilung ist unabhängig von den bedingenden Ereignissen.

- 3) Seien X und Y zwei verschiedene Verteilungsmengen (die z.B. zu unterschiedlichen Sozialprodukten gehören). Angenommen, für alle $x \in X$ und alle $y \in Y$ gelte $x P_i y$, dann gelte auch $(x_1 | h_1, \dots, x_k | h_k) R_i (y_1 | h_1, \dots, y_k | h_k)$ (sure thing postulate).

Annahme 3) beinhaltet, daß ceteris paribus p_i nicht einen Verteilungsprospekt, der für ihn weniger wünschenswerte Ergebnisse enthält, einem solchen mit günstigeren Ergebnissen vorzieht. Beide Postulate implizieren, daß p_i weder eine positive noch eine negative Einstellung zur Unsicherheit hat, die Entscheidungssituation selbst also keinen Nutzencharakter für ihn besitzt. Unter diesen Annahmen läßt sich nun folgender Satz beweisen:

Satz 3): Genügt eine individuelle Präferenzordnung R_i den Annahmen 1) - 3), dann ist sie mit einer kardinalen Nutzenfunktion U_i äquivalent, die jeder Menge L einen Wert zuordnet, der dem Erwartungswert des Nutzens der einzelnen Verteilungen $x \in X$ unter den Bedingungen h_1, \dots, h_k bei den Wahrscheinlichkeiten w_1, \dots, w_k entspricht, $U_i(L) = \sum_{j=1}^k w_j u_i(x_j)$ (vgl. Harsanyi 1976, 12):

Die Funktion $U_i(L)$ wird gewöhnlich als „v. Neumann-Morgenstern-

Nutzenfunktion“ bezeichnet und die Annahmen 1) - 3) werden auch „Bayesianische Rationalitätsprostate“ genannt.

Von diesem Konzept einer individuellen kardinalen Nutzenfunktion ausgehend, baut Harsanyi (bes. 1977 a+b) sein utilitaristisches Prinzip in zwei Schritten auf. Zunächst verweist er auf zwei grundverschiedene Situationstypen, in denen von Menschen Entscheidungen abverlangt werden. Im einen Situationstyp werden die Alternativen unter bloß persönlichen Aspekten betrachtet, hier urteilt und handelt der einzelne nur im Hinblick auf sein eigenes Wohlergehen und diese Beurteilung kommt in seiner individuellen Präferenzordnung zum Ausdruck. In dem anderen Situationstyp dagegen hat man nicht nur das zu beachten, was persönlichen Vorteil bringt, sondern darüberhinaus oder ausschließlich das, was moralisch geboten ist. Werden Handlungsfolgen unter moralischen Aspekten betrachtet, dann ist nicht mehr deren Nutzen für den Handelnden allein ausschlaggebend für die Beurteilung, sondern alle von der Handlung Betroffenen sind gleichermaßen in die Betrachtung mit einzubeziehen. Dies hält Harsanyi (1977a, 227) jedenfalls für eine notwendige Bedingung moralischen Urteilens. Solche Wertungen finden dann ihren Niederschlag in sozialen Präferenzen, die die Konstruktion einer besonderen Nutzenfunktion, der sogenannten individuellen Wohlfahrtsfunktion W_i der Person p_i erlauben. Eine individuelle Wohlfahrtsentscheidung soll nun vier Bedingungen gehorchen, die als Bedingungen einer rationalen moralischen Entscheidung angesehen werden können (ebd., 226; 1977b, 636f.).

- a) Die individuellen Präferenzen aller $p_i \in N$ folgen den Annahmen 1) - 3).
- b) Die sozialen Präferenzen aller $p_i \in N$ folgen ebenfalls den Annahmen 1) - 3).
- c) Angenommen, wenigstens für ein $p_j \in N$ und ein Paar $x_1, x_2 \in X$ gilt $x_1 P_j x_2$ und für kein $p_i \in N$ gilt $x_2 P_i x_1$, dann soll für jede individuelle Wohlfahrtsfunktion W_i gelten: $W_i(x_1) > W_i(x_2)$.
- d) Alle individuellen Nutzenfunktionen U_i haben einen gemeinsamen Maßstab, sind also vergleichbar.

Unter diesen Annahmen kann man den folgenden Satz beweisen:

Satz 4): Unter den Annahmen (a) - (d) hat die individuelle Wohlfahrtsfunktion W_i die Form $W_i = \sum_{i=1}^n a_i U_i$ mit $a_i > 0$ und $a_1 = \dots = a_n$.

Die individuelle Wohlfahrtsfunktion ist also die gewichtete Summe aller individuellen Nutzenfunktionen. Damit ist der entscheidende Schritt zur Formulierung des utilitaristischen Prinzips geleistet: Um ein gesellschaftliches Optimum zu erreichen, muß man sich jetzt für den Verteilungsprospekt entscheiden, für den die Summe der gleich-

gewichteten individuellen Nutzerwartungswerte maximal ist. Dabei erscheint jetzt der Gleichgewichtsgrundsatz als Folge der Annahme (d) und der Anforderungen an die Rationalität von Entscheidungen unter Unsicherheit und nicht als eigenes Prinzip, womit das Problem aber letzten Endes lediglich verschoben worden ist.

In einem zweiten Schritt versucht Harsanyi (1977a, 227 u. 1977b, 631ff) nun, den Zusammenhang zwischen individueller und moralischer Rationalität deutlicher zu machen. Dazu entwirft er folgendes Entscheidungsmodell (Equiprobability Model for Moral Value Judgments). Angenommen, ein $p_j \in N$ kennt zwar alle Nutzenfunktionen U_i der Mitglieder von N , ist aber nicht in der Lage, die richtige Zuordnung der U_i zu den entsprechenden Personen zu leisten, insbesondere kann er seine eigene Nutzenfunktion nicht identifizieren, d.h. p_j weiß nicht, welchen sozialen Platz er innerhalb von N einnimmt. Seine Informationen lassen nur den Schluß zu, daß er mit gleicher Wahrscheinlichkeit jeden der n Plätze einnehmen könnte. Für welche Verteilung würde sich in diesem Falle p_j nach den Grundsätzen Bayesianischer Rationalität entscheiden? Nun, nach Satz 3 muß er den Erwartungswert seines Nutzens maximieren, d.h. er muß $U_j = \frac{1}{n} \sum_{i=1}^n U_i$ bilden und also das $x \in X$ suchen, für das der arithmetische Durchschnitt der Nutzenerwartungswerte maximal ist. Dies ist aber auch genau das Ergebnis, das wir bei Anwendung von Satz 4 auf die Situation erhalten. D.h. also, daß in diesem Fall individuelle und moralische Rationalität zu dem gleichen Entscheidungsmechanismus führen.

An dem Ansatz von Harsanyi fordern im wesentlichen drei zentrale Punkte Kritik heraus: Sein Rationalitätskonzept, seine Forderung nach der Vergleichbarkeit der Nutzenfunktionen und die Funktion, die er dem „equiprobability model“ zuweist. Bezüglich seiner Rationalitätskriterien beteuert Harsanyi (vgl. 1977a, 224), daß die Annahmen 2 und 3 unmittelbar zwingenden Charakter hätten und sich eine weitere Diskussion erübrige. Daß dem nicht so ist, zeigt ein Blick in die Literatur, in der insbesondere das surething postulate heftig diskutiert wird (vgl. Diamond 1967; Sen 1970, chap. 7; Strasnick 1981; McClennen 1981). Aber ich will den Aspekt der rationalen Entscheidung unter Unsicherheit hier gar nicht weiter verfolgen, mir scheint das Rationalitätskonzept schon an einer viel elementareren Stelle seine entscheidende Schwäche zu besitzen: Die in Annahme 1) ausgedrückte Forderung nach einer vollständigen und transitiven Ordnung der Menge der Alternativen, der wir schon beim ordinalen Ansatz begegnet sind, beinhaltet eine Überforderung des Entscheidungsträgers und kann getreu dem Satze, daß nur das Mögliche auch machbar ist, nicht als sinnvolles Kriterium für Rationalität betrachtet werden. Wenn man nicht, wie gewöhnlich, lediglich den Ordnungsaspekt der Präferenzrelation im Auge hat, sondern auch auf den Nutzenaspekt

sieht, so wird man feststellen müssen, daß nur das „rational“ geordnet werden kann, was in seiner Unterschiedlichkeit auch Bedeutung für den Entscheidungsträger hat. Oder anders ausgedrückt, man kann nur in dem Maße eine Präferenzordnung erstellen, in dem den Alternativen ein Interesse entgegengebracht wird. Die Forderung nach Vollständigkeit der Präferenzordnung kann also nur dann sinnvoll erhoben werden, wenn unterstellt worden ist, daß man faktisch an allen sich bietenden Alternativen ein Interesse hat und damit ist sie von der Richtigkeit einer empirischen Behauptung abhängig. Sehen wir uns zur Verdeutlichung ein Beispiel an. Angenommen, jemand wird aufgefordert, seine Präferenz zwischen zwei Verteilungssituationen x_1 , x_2 darzulegen, wobei im Falle x_1 5000 DM p.a. und im Falle x_2 6000 DM p.a. auf ihn entfallen. Seine Antwort sei, daß ihn solche Fälle nicht interessieren, denn sollten die Verhältnisse so schlecht sein, daß er weniger als 24 000 DM p.a. als Einkommen hätte, dann würde er sowieso auswandern. Darf man diese Auskunft zu der Aussage uminterpretieren, der Betreffende sei zwischen x_1 , x_2 indifferent? Harsanyi selbst läßt nun eine „Fremdkorrektur“ von Präferenzen grundsätzlich zu: Wenn wir moralische Urteile abgeben, dann müssen wir die wohlverstandenen (true) Präferenzen der einzelnen betrachten, nicht ihre faktisch geäußerten. Damit ist gemeint, daß die faktischen Präferenzordnungen von „Fehlern“ bereinigt werden, die auf mangelhaften Informationen, ungenügenden Überlegungen oder emotionalen Affekten beruhen, ja Harsanyi (1977b, 646) geht sogar soweit, daß er rein antisoziale Präferenzen (also solche, die allein auf die Schädigung anderer abzielen) aus der faktischen Präferenzordnung eliminieren will. Aber diese Korrekturmöglichkeiten reichen ersichtlich nicht aus, um wenigstens die Vollständigkeit der wohlverstandenen Präferenzordnung zu sichern, denn durch den Fremdeingriff können nur Fehler in der Kalkulation korrigiert werden, nicht aber dürfen die Interessen selbst (mit der einen Ausnahme der antisozialen Präferenzen) tangiert werden. Dies verbietet der Grundsatz der Präferenzautonomie, den Harsanyi ansonsten auch akzeptiert.

Ähnlich problematisch ist die Forderung nach der durchgängigen Transitivität der Präferenzordnung. Wir sind es gewohnt, die Dinge unter verschiedenen Aspekten zu sehen und unter verschiedenen Bedingungen zu bewerten. Ich kann Rotwein (r) Bier (b) unter dem Aspekt der Bekömmlichkeit vorziehen ($r P b$) und Weißwein (w) dem Rotwein unter dem Aspekt des Geschmacks vorziehen ($w P r$), daraus auf $w P b$ zu schließen, kann nach gängigem Verständnis nicht als rational angesehen werden, denn genauso wäre die Präferenz $b P w$ denkbar und begründbar. Die Bedingtheit der Präferenzen führt dazu, daß wir für gewöhnlich zu einem Zeitpunkt über eine ganze Reihe von unterschiedlichen Präferenzordnungen verfügen und die

Frage der Rationalität entscheidet sich in unserem Leben in erster Linie daran, ob die Präferenzordnungen den Bedingungen angemessen sind oder nicht. Diese Problematik kann in der üblichen Nutzentheorie aber gar nicht artikuliert werden.

Einwände dieser Art lassen sich in gleicher Weise gegen das Rationalitätsverständnis im Rahmen der ordinalen Utilitarismusvariante vorbringen. Die Reduktion von Rationalität auf einige formale Konsistenzigenschaften stellt eine der Hauptschwächen des Neo-Utilitarismus dar, deren Bewältigung für den Aufbau eines zufriedenstellenden Ethikentwurfs eine vordringliche Aufgabe bildet.

Unter diesem Blickwinkel muß die apodiktisch vorgetragene Behauptung Harsanyis, in moralischen Zusammenhängen müsse die gleiche Form der Rationalität wie beim individuellen Handeln zur Anwendung gelangen und noch strenger beachtet werden (ebd., 637), erst recht Zweifel erwecken. Allerdings hat Strasnick (1981, 79f.) gezeigt, daß man den Satz (4) auch mit einer wesentlich schwächeren Annahme als der Annahme (b) erhalten kann, so daß sich deren Diskussion nicht lohnt und wir zur Diskussion der Annahme (d) übergehen können.

Die Kardinalität der individuellen Nutzenfunktionen gewährt allein noch nicht ihre Vergleichbarkeit und da sonst keine Eigenschaften der Nutzenfunktionen ersichtlich sind, die die Vergleichbarkeit sichern würden, ist diese Annahme im Hinblick auf die Ergebnisse notwendig. Meßtheoretisch gesehen beinhaltet dies, daß alle Nutzenfunktionen durch lineare Transformation ineinander überführbar sein müssen, eine reichlich kühne Forderung. Harsanyi (1977b, 638ff) argumentiert nun, daß, wie kühn die Forderung auch sein mag, uns im Falle moralischen Urteilens gar nichts anderes übrig bleibt, als so zu tun, als ob wir einen einheitlichen Maßstab hätten. Wir müssen uns in die Lage des anderen hineinversetzen und versuchen zu antizipieren, wie sich wohl bestimmte Handlungsfolgen für ihn auswirken würden. Wenn man die Betroffenen gut kennt, mag dies möglich sein, andernfalls ist die Antizipation nach bestem Wissen und Gewissen vorzunehmen.

Dies bedeutet natürlich, daß wir uns bei unseren Antizipationen auch irren und die eigentlichen Präferenzen des anderen völlig verfehlen können. Aber – und hierin ist Harsanyi sicher zuzustimmen – die Möglichkeit des Irrtums kann kein Grund dafür sein, daß wir uns ethisch-politischer Urteile enthalten. Die Forderung nach interpersonalem Nutzenvergleich kann nur heißen, unser Wissen und unsere Lebenserfahrung im Hinblick auf die Interessenlage anderer, so gut es geht, auszuwerten. Ich weiß nicht, inwieweit Harsanyi damit bewußt ein Kantisches Moment in die utilitaristische Ethik gebracht hat. Die „reine“ utilitaristische Lehre ist nämlich von einem strengen Folgenprinzip beherrscht: Gut ist eine Handlung dann, wenn ihre

Folgen tatsächlich nutzenmaximal sind. Nach den Ausführungen Harsanyis sollte man aber eine Handlung auch dann „gut“ nennen dürfen, wenn ihre Folgen nutzenmaximal sind oder zumindest die ihr zugrunde liegenden Nutzenabwägungen nach bestem Wissen und Gewissen getroffen wurden. In diesem modifizierten utilitaristischen Prinzip ist also zumindest subsidiär auch der Wille als Beurteilungsobjekt enthalten und so eine (schwache) Verbindung zur Kantischen Methode der Ethik geschaffen.

Was schließlich das equiprobability-model betrifft, so ist der Sinn dieses Modells prima facie nur schwer zu erkennen. Sicherlich wird damit keine realistische Entscheidungssituation beschrieben, das Modell ist ein Gedankenexperiment, eine bloße Denkmöglichkeit. Sollte dann dieses Geschichtchen nur die Funktion haben, exemplarisch darzulegen, was es heißt, von eigenen Belangen abzusehen und unparteilich zu urteilen? Man kann dies sicher so sehen, allerdings meine ich, daß doch noch ein tieferer Sinn hinter diesem Modell steckt.

Wie bereits oben im I. Abschnitt ausgeführt, ergibt sich für die Wohlfahrtsökonomie das Problem, wie man sich den Übergang von den individuellen zu den gesellschaftlichen Präferenzen vorzustellen hat. In der Ökonomie ist ja ganz unabhängig von allen Schulen die Vorstellung akzeptiert, daß der einzelne selbstbezogen versucht, seine individuellen Belange nach einer rationalen Zweck-Mittel-Abwägung durchzusetzen, mag auch im einzelnen umstritten sein, welche Form die handlungsleitende Rationalität besitzt. Diesem anerkannten Individualismus steht nun schroff die Forderung nach Maximierung der gesellschaftlichen Wohlfahrt gegenüber. Der Zustand des gesellschaftlichen Optimums im utilitaristischen Sinne stellt sich auch nicht wie etwa der pareto-optimale Zustand unter den Bedingungen vollständiger Konkurrenz automatisch ein, wenn die einzelnen selbstbezogen und rational handeln, das gesellschaftliche Nutzenmaximum läßt sich nur durch aktives, auf dieses Ziel ausgerichtetes Handeln erreichen. Man steht also vor einem Dilemma: Einerseits beurteilt man eine Gesellschaft, deren gesellschaftliche Wohlfahrt maximiert ist, als gut oder gerecht im weiteren Sinne, andererseits kennt man keine Argumente, um den einzelnen, dessen Selbstbezogenheit man ja akzeptiert, zum Streben nach gesellschaftlicher Wohlfahrt zu bewegen.

Die Ethik muß damit bloß appellativ bleiben, es sei denn, man könnte zeigen, daß die Erstrebung der gesellschaftlichen Wohlfahrt im eigenen Interesse liege, daß also das Handeln, welches üblicherweise als ethisch gerechtfertigt angesehen wird, schon unter dem Blickwinkel individueller Rationalität geboten ist. An solchen reduktionistischen Versuchen zur Ethikbegründung hat es in der Neuzeit nicht gefehlt. Wiederholt hat man versucht, das ethisch Gebotene als Gebot individueller Rationalität oder als Explikation des von der

menschlichen Natur Angelegten darzustellen. Diese Versuche sind als Fehlschläge anzusehen und dennoch scheint mir gerade der Vorschlag von Harsanyi – wie übrigens auch das Modell der „original position“ von Rawls (1972) – ein Zeichen dafür zu sein, daß zumindest abgeschwächte Modifikationen des reduktionistischen Programms heute noch lebendig sind. Sowohl Harsanyi als auch Rawls möchten zeigen, daß ein Zustand wenigstens denkbar ist, in dem die ethischen Grundsätze als Ausdruck der auf das Individuum bezogenen Rationalität erscheinen. Dies beweist, daß im Grunde genommen die ethischen Grundsätze nicht für sich allein als einsichtig gelten und man ihnen deshalb durch die Ankoppelung an die individuelle Rationalität ein zusätzliches argumentatives Gewicht verleihen möchte: Die Prinzipien der Rationalität und die Prinzipien der Ethik stehen sich, so hofft man gezeigt zu haben, nicht schroff gegenüber, sondern berühren sich zumindest im Bereich des Denkbaren.

2. Gerechtigkeit als Fairness

Charakteristisch für die betrachteten utilitaristischen Ansätze war, daß das Urteil über die Gerechtigkeit im weiteren Sinne einer Verteilung ganz unabhängig von den Einschätzungen der einzelnen bezüglich ihrer eigenen Position gefällt werden kann. Ein Zustand des gesellschaftlichen Wohlfahrts optimums ist gerecht im weiteren Sinne und er ist darüberhinaus auch gerecht im engeren Sinne, wenn zu seiner Ermittlung nur die individuellen Nutzenfunktionen herangezogen und in gleicher Weise gewichtet worden sind. Ob die einzelnen diesen Zustand auch für sich persönlich optimal ansehen, ist unerheblich.

In der Theorie der Fairness wird nun von einer gegenteiligen Betrachtungsweise ausgegangen. Der Grundgedanke dieser Theorie, die im wesentlichen auf D. Foley (1967) zurückgeht⁸, ist schnell geschildert. Betrachten wir wieder eine Verteilung $x \in X$. Mit x_i bezeichne ich das Güterbündel, das auf p_i entfällt, x_i ist also die i -te Spalte im Verteilungstableau x . Wir nehmen nun weiter an, daß p_i seine Situation mit der der anderen vergleicht, d.h. er bewertet die Anteile der anderen nach seiner Präferenzordnung. Gilt $x_i R_j x_j$, dann heißt dies, daß p_i das ihm zustehende Güterbündel x_i mindestens so schätzt wie das Güterbündel x_j von p_j und man sagt, p_i beneide p_j nicht. Eine Verteilung x , bei der für alle Paare $p_i, p_j \in N$ gilt $x_i R_j x_j$, wird dann „fair“ genannt. Das heißt, bei einer fairen Verteilung empfindet keiner auf die Position eines anderen Neid. Ein typisches Beispiel für eine solche faire Verteilung ist die Gleichverteilung, bei der jeder über das gleiche Güterbündel verfügt. Wir sehen also, daß in diesem

Fall das Urteil über den gesellschaftlichen Zustand von der subjektiven Sicht der Mitglieder von N abhängt: fair ist eine gesellschaftliche Ordnung, in der jeder mit seiner jeweiligen Position zufrieden ist.

So einfach sich dieses Konzept auch ausnimmt, man stößt sehr schnell auf Schwierigkeiten, wenn man es in ökonomischen Kontexten anwenden will. Wie sieht es z.B. mit dem Verhältnis von Fairness und Effizienz aus? Beim Neo-Utilitarismus in seiner kardinalen Variante gibt es im gesellschaftlichen Optimum trivialerweise keine Verschwendung, wenn auch umgekehrt nicht jeder pareto-optimale Zustand ein gesellschaftliches Optimum im utilitaristischen Sinne darstellt. Dagegen kann man sich sofort davon überzeugen, daß ein fairer Zustand durchaus nicht pareto-optimal sein muß. Angenommen, unter einer pareto-optimalen Verteilung x hätte p_i von einem Gut mehr als p_j und p_j würde p_i deshalb beneiden. Käme nun wegen Verschwendung ein Teil des Sozialproduktes nicht zur Verteilung mit der Folge, daß p_i unter diesen Umständen genausoviel besäße wie p_j , dann gäbe es zwar keinen Neid, dieser Zustand wäre aber auch nicht pareto-optimal. Die interessante Frage ist also, unter welchen Umständen ein pareto-optimaler Zustand auch fair ist. Als einen Schritt zur näheren Durchleuchtung dieses Problems sehen wir uns zunächst die Verhältnisse in einer reinen Tauschwirtschaft an. Es läßt sich hierfür folgender Satz beweisen.

Satz 5: Sei $\bar{x}_1 = \dots = \bar{x}_n$ die Anfangsausstattung der Mitglieder p_1, \dots, p_n einer Tauschwirtschaft (Gleichverteilung). Angenommen, alle p_i besitzen eine stetige und monotone Nutzenfunktion. Dann ist die zu einem Konkurrenzgleichgewicht gehörende Verteilung x^* pareto-optimal und fair. (Vgl. Varian 1974, 68)

Der Beweis dieses Satzes ist ganz einfach. Zunächst folgt die Pareto-Optimalität von x^* aus dem Hauptsatz der Wohlfahrtsökonomie. Im Konkurrenzgleichgewicht hat jeder p_i seine Nutzenfunktion relativ zu seiner Einkommensbeschränkung maximiert. Da wegen der Gleichverteilung der Anfangsausstattung jeder der gleichen Budgetbeschränkung unterliegt, ist es ausgeschlossen, daß ein p_i bei einem Güterbündel landet, das er weniger schätzt als das eines anderen (er hätte dieses ja ebensogut kaufen können). Also gilt für alle Paare p_i, p_j : $x_i^* R_i x_j^*$, d.h. die Verteilung x^* ist fair. Dieses Ergebnis ist wegen seiner äußerst rigiden Annahmen wenig zufriedenstellend und die Schwierigkeiten wachsen schnell an, wenn man auch nur ein Stückchen auf die Realität zugehen möchte.

Angenommen, man betrachtet eine Ökonomie mit Produktion, deren Mitglieder nicht identisch bezüglich ihrer produktiven Fähigkeiten sind, so daß ihr Beitrag zum Sozialprodukt unterschiedlich ausfällt. Was soll unter diesen Bedingungen als „gleiche Anfangsaus-

stattung“ betrachtet werden? Gleiches Einkommen kann nicht gefordert werden, denn im Konkurrenzgleichgewicht (also in einem pareto-optimalen Zustand) leitet sich der Lohn von der Grenzproduktivität der Arbeit ab und diese wiederum ist je nach Geschick und Fähigkeit verschieden. Es ist also nicht verwunderlich, wenn unter den pareto-optimalen Verteilungen einer Ökonomie mit Produktion, die den Standard-Bedingungen gehorcht, keine faire Verteilung zu finden ist.

Aus dieser Situation hat man versucht, auf zwei verschiedene Weisen einen Ausweg zu finden. Den einen haben Varian, Schmeidler, Pazner u.a. beschritten, indem sie den Fairness-Begriff modifiziert haben. So hat Varian die Begriffe „wealth-fair“ und „income-fair“ für Ökonomien mit Produktion geprägt. Eine Verteilung wird demnach „wealth-fair“⁹ genannt, wenn kein Individuum p_i dem eigenen Bündel aus nachgefragten Konsumgütern und selbstgeleistetem Output ein derartiges Bündel eines anderen Individuums vorzieht. D.h. also, daß überhaupt nur solche Neidäußerungen ernsthaft betrachtet werden, bei denen der Anspruch auf ein anderes Konsumgüterbündel mit der Bereitschaft verbunden ist, auch den Output des anderen zu erbringen: Nur der darf auf das Einkommen eines Karajan „mit Recht“ neidisch sein, der auch so gut dirigieren kann und bereit ist, das entsprechende Arbeitspensum zu leisten.

Mit „income-fair“ (vgl. Varian, 1974, 75) wird eine Verteilung bezeichnet, der gemäß ein unter den Bedingungen der vollständigen Konkurrenz erwirtschaftetes Sozialprodukt so aufgeteilt wird, daß jeder ein Bündel aus Konsumgütern und Freizeit von gleichem Wert erhält. Jeder hat dann gleiches Einkommen, nur müssen die einzelnen je nach Geschick und Fähigkeit unterschiedlich lang dafür arbeiten.

Und schließlich sei noch der von E. Pazner und D. Schmeidler (1978) geprägte Begriff der „egalitarian-equivalence“ erwähnt¹⁰: Eine Verteilung x ist „egalitarian-equivalent“, wenn unter den Bedingungen einer hypothetischen „income-fairen“ Verteilung kein p_i mehr erhalten würde. Das heißt also, daß x selbst nicht fair sein muß, nur könnte es unter fairen Bedingungen keinem besser gehen.

Arbeitet man mit diesen Fairness-Begriffen, dann existieren in der Tat pareto-optimale Zustände, die fair im jeweiligen Sinne sind. Ich möchte diese Ansätze im einzelnen nicht diskutieren, weil sie in meinen Augen recht offensichtliche ad-hoc-Konstruktionen sind, die so ausgedacht wurden, daß sie mit dem Begriff der Pareto-Optimalität verträglich sind und gerade noch etwas mit dem ursprünglichen Fairness-Begriff zu tun haben.

Zwei interessante Punkte müssen wir jedoch festhalten. Bislang wurde ein Konzept von Gerechtigkeit im engeren Sinne verwendet, das darauf hinauslief, jeden in gleicher Weise und nur nach seinen Bedürfnissen zu beachten. Mit den modifizierten Fairness-Begriffen

wird zum ersten Mal das Verhältnis von Anspruch auf Bedürfnisbefriedigung und Leistungseinsatz bzw. Bereitschaft zum Leistungseinsatz wenn auch nicht diskutiert, so doch wenigstens artikuliert. Darüberhinaus gibt der Pazner-Schmeidler'sche Ansatz noch Anlaß zu einer interessanten Frage, die uns später noch beschäftigen wird: Angenommen, den unter einer ungerechten Verteilung Benachteiligten ginge es unter gerechten Bedingungen nicht besser: Welchen Sinn sollte dann eine Forderung nach mehr Gerechtigkeit haben?

Der zweite Versuch, einen Ausweg aus den Schwierigkeiten zu finden, geht auf Gärdenfors (1978) zurück. Gärdenfors verwendet das Fairness-Konzept, um eine Wertung unter pareto-optimalen Verteilungen zu begründen. Ausgangspunkt für seine Überlegungen ist dabei der folgende Satz:

Satz 6: Ist eine Verteilung x pareto-optimal, dann gibt es mindestens einen p_j , der von niemandem beneidet wird und mindestens einen p_j , der niemanden beneidet.

Zum Beweis dieses Satzes nehmen wir zunächst an, jeder würde von anderen beneidet. Dann gäbe es eine Teilmenge $M \subset N$ derart, daß ihre Mitglieder einen „Neid-Zirkel“ bilden: p_1 beneidet p_2 , p_2 beneidet p_3 , . . . , p_m beneidet p_1 . Man könnte sich dann eine Verteilung x' denken, bei der für alle Mitglieder aus $N \setminus M$ die Verhältnisse unverändert bleiben, in M aber jeder an die Stelle rückt, die er der eigenen vorzieht. Diese Verteilung x' würde von allen $p_j \in N$ mindestens so gut geschätzt werden wie x , die Mitglieder aus M würden x' sogar x vorziehen. Dies widerspricht jedoch der Annahme, daß x pareto-optimal ist. Entsprechend läßt sich der Fall beweisen, daß es mindestens einen p_j gibt, der niemanden beneidet.

Diesen Gedanken benutzt Gärdenfors zur Konstruktion eines ordinalen Fairness-Maßes für pareto-optimale Verteilungen. Dazu definiert er (ebd., 68f.): Eine Verteilung x heißt fairer als x' , wenn gilt:

a) Alle, die unter x Neid empfinden, empfinden ihn auch unter x' . Es gibt jedoch mindestens einen $p_j \in N$, der zwar unter x' Neid empfindet, nicht jedoch unter x . b) Sollte ein p_j unter x schlechter gestellt sein als unter x' , so nur deshalb, weil es einen p_j gibt, der unter x' schlechter gestellt war als unter x und der in jedem Falle schlechter gestellt ist als p_j .

Es bedeutet b) insbesondere unter Berücksichtigung von a), daß der Gedanke der Fairness eine Umverteilung von den Reichen zu den Ärmeren zuläßt, allerdings dürfen die Ärmeren die Reichen nicht „überholen“. Gärdenfors kann nun zeigen, daß die Relation „fairer als“ transitiv und irreflexiv ist und eine partielle Ordnung pareto-optimaler Zustände ermöglicht.

Wenn wir Gärdenfors' Vorschlag mit den anderen Vorschlägen vergleichen, wird deutlich, daß er ein neues Element in die Diskussion gebracht hat. Varian, Pazner, Schmeidler u.a. verwenden das Fairness-Konzept, um nach einem schlichten Ja-Nein-Schema ein Urteil über eine bestimmte Ordnung zu ermöglichen: Ist ein bestimmter relevanter ökonomischer Zustand fair? Wenn ja, ist alles in Ordnung, wenn nein, muß u.U. solange an dem Fairness-Begriff gezerrt werden, bis schließlich ein positives Urteil herauskommt. Diesen Hang zum undifferenzierten Ja-Nein-Urteil finden wir übrigens auch bei den Neo-Utilitaristen. Auch sie wollen letztlich nur die Frage entscheiden, ob in einem bestimmten Zustand optimale gesellschaftliche Wohlfahrt herrscht oder nicht. Dabei wird völlig übersehen, daß normative Vorstellungen im Sinne von Ideen durchaus handlungsleitende Funktion haben können, auch wenn sie sich als solche nicht in vollem Umfang realisieren lassen. Und genau dieser Punkt wird mit dem Beitrag von Gärdenfors deutlich gemacht. Ein Konzept von Fairness liefert – wie prinzipiell jedes andere Konzept von Gerechtigkeit im weiteren oder engeren Sinn – dann für eine bestimmte Verteilungspolitik eine argumentative Stütze, wenn sich zu jedem Zustand ein in eindeutiger Weise „besserer“ angeben läßt. Solche Ideen haben also ihre Funktion als Kriterien für die Festlegung einer politisch-ethischen Richtung, sie sind keine Beschreibungen potentieller gesellschaftlicher Zustände.

3. Zwischenbilanz

Wenn wir an dieser Stelle die skizzierten Ansätze noch einmal auf die in ihnen zum Ausdruck kommenden Gerechtigkeitsprinzipien hin betrachten, dann lassen sich im wesentlichen vier Punkte kritisch vermerken.

1) Auffallend ist zunächst, daß das Gebot der Wohlfahrtsmaximierung im Neo-Utilitarismus grundsätzlich an institutionelles Handeln gebunden ist. Das Individuum darf in einer sehr selbstbezogenen Weise nach seinem individuellen Glück streben, seine private Nutzenfunktion (Präferenzordnung) wird kaum zur Diskussion gestellt. Lediglich wenn es um den institutionellen Interessenausgleich im Rahmen einer Verfassung geht (so beim Programm der Social Choice Theory), bzw. wenn der einzelne in Verantwortung vor der Öffentlichkeit zu handeln hat (so bei Harsanyi), spielt das Nützlichkeitsprinzip eine Rolle.

2) Für den Neo-Utilitarismus ist gerecht im weiteren Sinne eine Gesellschaft, die ein höchstmögliches Maß an Bedürfnisbefriedigung ge-

währt. In einer solchen Gesellschaft darf es große Unterschiede bezüglich des Wohlfahrtsniveaus der einzelnen geben und entsprechend groß mag das Potential an Unzufriedenheit in ihr sein, jedoch werden zumindest einige ihren Vorstellungen von einer wünschenswerten Lebensführung recht nahe kommen können.

Dies bedeutet indes nicht, daß diese Gesellschaft auch gerecht insoweit wäre, als jeder nach seinen Bedürfnissen gewichtet würde, wie es nach einer Interpretation des „Hauptsatzes“ der distributiven Gerechtigkeit zu geschehen hätte. Die individuellen Bedürfnisse kommen gesellschaftlich über ein Aggregationsverfahren zur Geltung, bei dem jeder mit dem gleichen Gewicht berücksichtigt wird. D.h., daß auch die Bedürftigsten in einem qualitativen Sinne keine Auszeichnung erfahren. Gerechtigkeit im engeren Sinne kommt im Neo-Utilitarismus nur als prozedurale Gleichheit zum Tragen: Als einziges Qualifikationsmerkmal der Menschen wird ihre Eigenschaft, Träger einer Nutzenfunktion bzw. Präferenzordnung zu sein, in Betracht gezogen. Da die Menschen bezüglich dieses Merkmals gleich sind, müssen sie konsequenterweise gleichgewichtig in die ethische Beurteilung einbezogen werden. Gerechtigkeit wird so in ihrem engst möglichen Sinne berücksichtigt.

3) Im Rahmen des Fairness-Ansatzes ist gerecht im weiteren Sinne ein „Minimal-Zustand“: Ob man seinen eigenen Vorstellungen entsprechend leben kann, ist nicht so wichtig, entscheidend ist nur, daß niemand besser leben kann als man selbst. Nun ist weder die Theorie der Fairness noch der Neo-Utilitarismus durch eine Theorie der unantastbaren Rechte des Menschen gestützt, so daß beide Ansätze Rechtfertigungen für Umverteilungsmaßnahmen bieten, die den status quo des einzelnen u.U. gravierend beeinträchtigen können. Während es aber bei einer utilitaristischen Umverteilung immer einen Nutznießer gibt, ist nach der Fairness-Theorie ein Zustand denkbar, in dem Gerechtigkeit im weiteren Sinne durch schlichte Konfiskation herbeigeführt wird: Gerechtigkeit als Nivellierung nach unten. Daß sich hierbei Gerechtigkeit als Fairness leicht mit dem Bild einer Gesellschaft von gleichermaßen trost- und hoffnungslosen Menschen verbinden läßt, scheint mir der entscheidende Mangel dieser Theorie zu sein. So kann die dem ursprünglichen Fairnessansatz innewohnende Tendenz zur materialen Gleichheit auch nicht als Ausdruck von Gerechtigkeit im engeren Sinne interpretiert werden, denn niemand ist verpflichtet, seinen Neidäußerungen den Hauptsatz der distributiven Gerechtigkeit zugrunde zu legen, so daß an diesem Maßstab gemessen, die interpersonellen Vergleichsurteile völlig unbegründet sein können. Letztlich kommt Gerechtigkeit auch hier nur als prozedurale Gleichheit ins Spiel: jede Neidäußerung ist in gleicher Weise zu berücksichtigen. Bei den modifizierten Fairness-Ansätzen wird

dann zwar das ursprüngliche Konzept mit zusätzlichen ethischen Forderungen angereichert, welche die Mitglieder einer Gesellschaft bei ihren Neidäußerungen zu beachten haben. Aber der faire Zustand ist dann eben kein Zustand einer voraussetzungslosen Zufriedenheit mehr, sondern einer, in dem man die Ergebnisse akzeptiert, die durch die Anwendung einer Ethik zustande gekommen sind, der man sich selbst unterworfen hat. Ich hoffe, damit ist deutlich geworden, daß die Theorie der Fairness kein erfolversprechender Ansatz ist, den Schwierigkeiten des Neo-Utilitarismus aus dem Weg zu gehen.

4) Daß die Gerechtigkeitskonzeptionen in der Wohlfahrtsökonomie so sparsam ausfallen, hat zumindest indirekt etwas mit dem Grundsatz der Präferenzautonomie zu tun, der allen Ansätzen gemeinsam ist.

Die subjektiven Wünsche, Ansprüche und Neidäußerungen sind so zu akzeptieren, wie sie faktisch geäußert werden, lediglich bezüglich der Korrektur von Irrtümern und kalkulatorischen Fehlern wird manchmal eine Ausnahme gestattet.¹¹ Da man also nicht bereit ist, die subjektiven Äußerungen einer kritischen Würdigung zu unterwerfen, begnügt man sich – trotz aller formaler Präzision – mit einem inhaltlich reichlich schwammigen Nutzen- bzw. Präferenzbegriff. Üblicherweise führt man ja in der Wohlfahrtsökonomie den Nutzenbegriff über den Präferenzbegriff ein. Zieht p ein Güterbündel x einem Güterbündel x' vor, so darf man sagen, daß x einen größeren Nutzen für p hat als x' . Diese semantische Normierung verweist aber nicht auf einen sonstwie begründeten Zusammenhang zwischen Präferenz und Nutzen. Ich habe oben der Einfachheit halber von Nutzen im Sinne von „Beitrag zur Bedürfnisbefriedigung“ gesprochen. Dies ist wohl den meisten Ökonomen schon zu substanziiert, A. Sen (1979, 463) spricht z.B. von „utility“ nur als „a person's conception of his own wellbeing“ und selbst eine solche Formulierung ist für Mirrlees (1982, 64 u. 83) noch zu eng.¹² Generell ist man wohl der Ansicht, daß keine besonderen Gründe dafür angeführt werden müssen, um etwas als „nützlich“ zu beurteilen, es reicht, daß man es möglichen Alternativen vorzieht, aus welchen Gründen auch immer. Während aber im Neo-Utilitarismus wenigstens sporadisch noch eine Tendenz zu inhaltlicher Rationalität ausgemacht werden kann, scheidet auch diesbezüglich die Theorie der Fairness schlechter ab, läßt sie doch mit der Betrachtung von Neidäußerungen explizit Irrationalitäten und emotionale Entgleisungen als Beurteilungsgrundlage zu.

Die Abstraktion von den Gründen für bestimmte Präferenzen schließt somit auch die Beachtung von Unterschieden in Bildung, Erziehung, Ausbildung, Intelligenz und Lebensart der Menschen aus. So werden die Menschen nicht als gleich in bezug auf ihre Fähigkeiten

zu lernen, wie autonom zu entscheiden ist, betrachtet, es wird vielmehr unterstellt, sie wären gleichermaßen faktisch autonome Subjekte. Diese Unterstellung, die häufig völlig unreflektiert in die theoretischen Ansätze einfließt, verhindert letztlich eine gehaltvollere Behandlung des Gerechtigkeitsproblems.

4. Gerechtigkeit bei J. Bentham und J. St. Mill

4.1. J. Bentham

Die Gerechtigkeitskonzeption des Neo-Utilitarismus hat ihre Wurzeln in einer spezifischen Richtung des klassischen Utilitarismus, dem Benthamianismus, der vor allem von John Austin und James Mill propagiert wurde. Bentham selbst widmete dem Gerechtigkeitsproblem kaum mehr als beiläufige Aufmerksamkeit. Aus den wenigen Stellen, an denen er explizit darauf eingeht, läßt sich jedoch erschließen, daß er den Begriff „Gerechtigkeit“ in zweifacher Bedeutung verwendet.

In der einen Bedeutung ist Gerechtigkeit allein im Rechtswesen angesiedelt und meint im wesentlichen „Unparteilichkeit“. Das Musterbeispiel ist für ihn dabei der gerechte Richter, der die Gesetze ohne Ansehen der Person anwendet (Bentham, 1823, 113f.), wobei zunächst unerheblich ist, welchen ethischen Status die Gesetze selbst besitzen.

In der anderen Bedeutung bezieht Bentham den Begriff „Gerechtigkeit“ auch auf das Verfahren der Gesetzgebung selbst. Ein gesellschaftlicher Zustand, der dem Nutzenprinzip entspricht, besitzt „natural justice“ (ebd., 17f.) und so ist der Gesetzgeber aufgefordert, „gerechte“, d.h. also nutzenmaximale Gesetze zu erlassen. Andere Gerechtigkeitsauffassungen will Bentham nicht zulassen, sie dienen seiner Meinung nach nur der intellektuellen Bemäntelung neidkittierter Angriffe auf das Nutzenprinzip. Aufschlußreich ist hier eine Anmerkung (ebd., 125f.), in der Bentham schreibt: „Sometimes, in order the better to conceal the cheat (from their own eyes doubtless as well as from others) they set up a phantom of their own, which they call Justice: whose dictates are to modify (which being explained, means to oppose) the dictates of benevolence. But justice, in the only sense in which it has a meaning, is an imaginary personage, feigned for the convenience of discourse, whose dictates are the dictates of utility, applied to certain particular cases“.

Es ist nicht einfach, zwischen diesen beiden Verwendungen des Wortes „Gerechtigkeit“ eine Verbindung herzustellen. H. Bedau (1963, 292) hat in einer Analyse der Gerechtigkeitstheorie Ben-

thams folgenden Interpretationsvorschlag gemacht, der den Intentionen von Bentham wohl am nächsten kommen dürfte. Eine moralische oder gesetzliche Norm wird gerecht genannt, wenn ihre Befolgung in einer bestimmten Klasse von Situationen zu einem nutzenmaximalen Ergebnis führt; eine Handlung wird gerecht genannt, wenn sie in der unparteilichen Befolgung einer Norm besteht oder relativ zu einer bestimmten Situation zu einem nutzenmaximalen Ergebnis führt.

Daß diese terminologischen Bestimmungen mehr als unglücklich sind, liegt auf der Hand. Um gerecht im Sinne von unparteilich zu handeln, ist es nicht erforderlich, daß die Norm, der man folgt, selbst gerecht ist. Damit wird die Möglichkeit eröffnet, daß eine unparteiliche Handlung gegen das Gebot der Nutzenmaximierung verstößt und also als ungerecht einzustufen ist. Bentham liefert hierfür keinen Lösungsvorschlag, ja es spricht viel dafür, daß er dieses Problem gar nicht erkannt hat. Deutlicher wird hingegen sein Mitstreiter James Mill. Er betont explizit, daß im Zweifelsfall die nutzenmaximierende Handlung den Vorzug vor der gesetzestreuenden bekommen soll. Damit bezieht er eine eindeutig handlungs-utilitaristische Position (1833, 163). Normen haben für ihn den Charakter von „rules of thumb“, d.h. sie können nur für die Situationstypen Verbindlichkeit beanspruchen, die bei ihrer Aufstellung in Betracht gezogen worden sind, ansonsten ist ihre Geltung vorbehaltlich, womit der Begriff „Gerechtigkeit“ nur noch in einer Bedeutungsweise verwendet wird.

Nun finden wir bei Bentham neben dem Nützlichkeitsprinzip noch ein anderes Prinzip, nämlich sein berühmtes: „Each count for one, none for more than one.“¹³ Bentham bringt diesen Satz nicht mit dem Begriff der Gerechtigkeit in Verbindung, es scheint mir aber dennoch nicht angebracht zu sein, ihn als ein dem Nutzenprinzip untergeordnetes materiales Gleichheitsprinzip zu interpretieren, so wie dies Bedau (1963, 294) andeutet. Das Prinzip wird bei Bentham im Zusammenhang mit einer Verfassungsdiskussion aufgestellt und soll das Gewicht des einzelnen Bürgers in bezug auf die Gesetzgebung festlegen. Von daher gesehen, ist es wohl am ehesten als Forderung der Unparteilichkeit bei der Ermittlung des Gesamtnutzens zu interpretieren. Als solches ist es nicht durch das Nutzenprinzip relativierbar und entspricht genau dem, was wir oben „Prinzip der prozeduralen Gleichheit“ genannt haben. In diesem Sinne versteht übrigens auch J. St. Mill diesen Satz (1861, 257). Mill meint zwar, dieses Prinzip sei eine analytische Folgerung aus dem Nützlichkeitsprinzip, seine vorsichtige Formulierung dieser Behauptung läßt jedoch darauf schließen, daß er diesbezüglich vielleicht doch seine leisen Zweifel hatte (ebd., 258).

4.2. J. St. Mill

Wie bei allen Utilitaristen, so ist auch bei J. St. Mill die Gerechtigkeit von Anfang an in einen gesellschaftlichen Kontext eingebettet. Allerdings wird erst bei Mill deutlich, daß dies nicht nur eine zufällige Sache der Darstellung oder der pragmatischen Einstellung ist, sondern daß die Rede von Gerechtigkeit eine geordnete Gesellschaft notwendig voraussetzt. Um Gerechtigkeit kann nach Mill nur dort gerungen werden, wo die Mitglieder einer Gesellschaft Rechte besitzen und ihnen Pflichten auferlegt sind. Ein Recht besitzen, bedeutet dabei, einen Anspruch gegen die Gesellschaft haben, d.h. der Rechtsinhaber verfügt über etwas, das die anderen Mitglieder der Gesellschaft respektieren und schützen sollen: „If he has what we consider a sufficient claim, on whatever account, to have something guaranteed to him by society, we say that he has a right to it,“ (Ebd., 250; vgl. a. Lyons 1978 u. Gaus 1980). Jedem Recht einer Person korrespondiert die entsprechende Pflicht der anderen Mitglieder der Gesellschaft, diese Position zu wahren. Umgekehrt gilt aber nicht, daß jeder Pflicht auch ein entsprechendes Recht korrespondiert, womit Mill die alte Unterscheidung von vollkommenen und unvollkommenen Pflichten aufnimmt (ebd., 247). Ein typisches Beispiel für unvollkommene Pflichten sind Barmherzigkeit und Wohltätigkeit, also Pflichten, zu deren Ausübung man zwar aufgerufen ist, auf deren Zuewendung aber niemand einen Rechtsanspruch besitzt.

Die Rechtspositionen der Menschen stehen nun aber nicht unvermittelt und gleichberechtigt nebeneinander. Sie besitzen durchaus verschiedene Wertigkeiten und können demnach hierarchisch geordnet werden. D.h. im Konfliktfall kann die Verletzung eines Rechts durch Verweis auf ein höherrangiges Recht erlaubt, ja sogar geboten sein. So mag es zur Rettung eines Menschenleben durchaus geboten sein, daß Eigentumsrechte verletzt werden (S. 259).

Als unrecht werden von Mill solche Handlungen angesehen, die die Rechtspositionen einer Person in ungerechtfertigter Weise verletzen und gerecht wird eine Handlung genannt, wenn ihre Unterlassung unrecht wäre und ihre Ausführungen eine aus einer Rechtsposition heraus einforderbare Pflicht (S. 247). Als äußeres Zeichen für die Unrechtmäßigkeit einer Handlung sind die gegen sie verhängten Sanktionen anzusehen, und so können wir eine Gesellschaft gerecht nennen, in der alle für unrecht erkannten Handlungen Sanktionen ausgesetzt sind.

Es sieht also prima facie so aus, als würde Mill zwei ethische Prinzipien verwenden. Einmal ein Gerechtigkeitsprinzip, das die Unantastbarkeit von Rechten garantiert und das zur Bestimmung von un-rechten Handlungen erforderlich ist (denn Handlungen werden bei Mill nicht einfach als unrecht qualifiziert, wenn sie gegen das Nüt-

lichkeitsprinzip verstoßen!) und zum anderen das Nützlichkeitsprinzip, dem zufolge unter Wahrung der erworbenen Rechte die gesellschaftliche Wohlfahrt zu maximieren ist. Es ist deshalb einigermaßen verblüffend, wenn Mill sagt, es sei evident, daß alle Fälle von Gerechtigkeit auch Fälle von Nützlichkeit seien (S. 259) und die Gerechtigkeit auf Nützlichkeit gegründet sei (S. 255).

Nun führt Mill zwar aus, daß das, was den Menschen als Recht zustehe, nur über das Nützlichkeitsprinzip bestimmt werden könne: „To have a right, then, is, I conceive, to have something which society ought to defend me in the possession of. If the objector goes on to ask why it ought, I can give him no other reason than general utility.“ (S. 250) Des weiteren wird die Hierarchisierung der Rechte ebenfalls über das Nützlichkeitsprinzip begründet: Nur ihm entsprechend wird entschieden, welche Rechtsposition im Konfliktfall gerechtfertigt verletzt werden darf (S. 259). Hieraus folgt indes noch nicht unmittelbar, daß ein Prinzip, welches ungerechtfertigte Eingriffe in die Rechte von Personen verbietet, selbst Ausfluß des Nützlichkeitsprinzips ist. Denn verbietet es etwa allein das Nützlichkeitsprinzip, daß man jemandem sein wohlverdientes Einkommen wegnimmt, um es einer Person zuzuführen, die damit ein höheres Nutz-niveau erreichen kann? Mill versucht seine Behauptung über die Annahme eines speziellen Sicherheitsbedürfnisses zu begründen. Dieses Bedürfnis, das in den Augen Mills unmittelbar nach dem Nahrungsbedürfnis kommt, soll bei allen Menschen anzutreffen sein und somit den Status einer anthropologischen Konstante haben (S. 251). Es ist auf die Abwehr von Eingriffen gerichtet und weil es – modern gesprochen – auf jeder individuellen Nutzenskala so weit oben zu finden ist, darf eine nutzenmaximierende gesellschaftliche Ordnung willkürliche Akte nicht zulassen. In diesem Sinn soll also die Forderung nach Gerechtigkeit auf das Nützlichkeitsprinzip gegründet sein. Auf unser gerade erwähntes Beispiel angewandt, bedeutet dies, daß nicht nur der Nutzen einer Einkommensumverteilung in Anschlag gebracht werden darf, sondern auch die Nutzeneinbuße, die in dem Eingriff und der damit verbundenen Gefährdung der Rechtsicherheit liegt, berücksichtigt werden muß.

Mit dieser Argumentation ist Mill allerdings ein Opfer seiner eigenen psychologisierenden Redeweise geworden. „Sicherheit“ als Bedürfnis zu bezeichnen und in einem Atemzug mit dem Nahrungsbedürfnis zu nennen, erweist sich als ausgesprochen unglücklich und paßt bei näherem Hinsehen auch nicht in die Mill'sche Systematik. Unter dem Nutzenaspekt werden nämlich alle Wünsche und Bedürfnisse betrachtet, deren Erfüllung uns Befriedigung gewähren. Was immer dazu auch gehören mag, Sicherheit kann jedenfalls in einen solchen Katalog nicht aufgenommen werden. Denn einmal kann der Zustand der Rechtssicherheit im Rahmen des individualistischen

Handlungsverständnisses nicht als Mittel für bestimmte Zwecke begriffen werden, da die Realisierung dieses Zustandes für den einzelnen gar nicht im Bereich des Machbaren liegt. Und zum anderen kann Sicherheit auch nicht als eigenständiger Zweck angesehen werden. Es ist zwar richtig, daß Unsicherheit auch psychische und physische Belastung mit sich bringen kann, aber eben nur mittelbar dadurch, daß wir etwas bewahren wollen oder anstreben, was der Bedrohung ausgesetzt ist: Wer kein Ziel anstrebt, braucht auch keine Sicherheit. Sicherheit ist also eine gesellschaftliche Bedingung dafür, daß unsere Lebensplanung in die Realität umgesetzt werden kann und ihre Gewährleistung macht das Anstreben von Zielen überhaupt erst erfolgversprechend. Wären wir einer ständigen Bedrohung ausgesetzt, dann könnten wir nur dem Moment leben, eine über den Augenblick hinausgehende Planung und Lebensgestaltung hätte unter diesen Bedingungen keinen Sinn. Es sei übrigens angemerkt, daß auch im Neo-Utilitarismus à la Harsanyi, in dessen Rahmen man Unsicherheit im statistischen Sinne explizit berücksichtigt, Rechtssicherheit als selbstverständlich vorausgesetzt wird, da die Beziehung zwischen bedingendem Ereignis und seinem Ergebnis als sicher gilt: Ein Roulette-spiel, bei dem im Falle des Gewinns die Auszahlung im Gutdünken der Bank liegt, ist schlechterdings unkalkulierbar. Halten wir also fest: Sicherheit ist eine notwendige Bedingung für die Durchführung einer vernünftigen Nutzenkalkulation, sie kann deshalb selbst keinen Nutzwert haben. Damit schließt sich aber der Kreis nicht so, wie es sich Mill gedacht hat und es ist ihm letztlich nicht gelungen, das Prinzip der Gerechtigkeit in den Utilitarismus zu integrieren. Gleichwohl ist bei ihm ein Weg angelegt, wie dies geschehen könnte und ich möchte in einem letzten Abschnitt diesem Gedanken noch etwas nachgehen. Zu einem besseren Verständnis (wenn auch nur als grobe Orientierung) muß ich jedoch noch einen kleinen Exkurs zum Begriff des Glücks im Utilitarismus vorausschicken (als guten Überblick dazu vgl. Höffe 1979, 120-159).

5. Kleiner Exkurs zum Glücksbegriff im klassischen Utilitarismus

Die Theorie des Glücks bei Bentham (1823) ist kurz, prägnant und wenig differenziert: Dem Glück (happiness) förderlich ist jede Handlung, die Lust (pleasure) hervorruft und/oder Schmerz (pain) verhindert. Was uns im einzelnen Lust und Schmerz verschafft, listet Bentham in einem umfangreichen Katalog auf (ebd., Chap. V). Aus diesem Katalog wird deutlich, daß er einen weitgehend neutralen Lustbegriff verwendet. Lust kann gleichermaßen aktivem Tun und passiver Beschaulichkeit entspringen, auch gibt es keine Tätigkeiten, Din-

ge oder Umstände, die von sich aus geeignet wären, in unterschiedlichem Maße Lust oder Schmerz zu verschaffen, d.h. der Glücksbegriff ist letztlich unabhängig von kulturellen Wertungen. Damit redet Bentham keinem Vulgär-Utilitarismus das Wort — wird sein Ausspruch „quantity of pleasure being equal, push-pin is as good as poetry“¹⁴ manchmal auch in diesem Sinn verstanden —, dies ist vielmehr nur Ausdruck des Grundsatzes der Präferenzautonomie: was einem Menschen Lust oder Schmerz bereitet, kann dieser nur allein beurteilen.

Das Streben nach Lust und die Aversion gegen Schmerz sind für Bentham anthropologische Fakten, „Nature has placed mankind under the governance of two sovereign masters, pain and pleasure,“ (1823, 1). Allerdings begründet er die Forderung nach Nutzenmaximierung nicht unmittelbar aus diesem Faktum. Er fordert vielmehr dazu auf, das bewußt und vernünftig zu betreiben, was sonst nur un deutlich gefühlt und artikuliert wird. Jeder Mensch möge sich anhand einer Nutzenbilanz für oder gegen eine Handlung entscheiden und nicht aufgrund spontaner Regungen und Vermutungen. Auch die Aufforderung an den politisch verantwortlich Handelnden, das Glück derjenigen, deren Interessen zur Diskussion stehen, zu maximieren, wird nicht über die Anthropologie „bewiesen“.¹⁵

Selbst wenn man die psychologistische Lust-Schmerz-Redeweise von Bentham weitherzig und nicht im Sinne moderner Verhaltenspsychologie interpretiert, bleibt jedenfalls festzuhalten, daß sich für ihn Glück als unmittelbare Folge von Handlungen einstellt. Diese direkte Bewertung von Handlungsfolgen danach, ob sie die Erfüllung von Wünschen bzw. die Befriedigung von Bedürfnissen bringen, ist ein Punkt mehr, den der Neo-Utilitarismus mit Bentham gemeinsam hat, wenn dann auch bezüglich der Nutzenkalkulation getrennte Wege gegangen werden.

Bei J.St. Mill ist die Theorie des Glücks viel komplexer und differenzierter, an manchen Stellen auch dunkel und widersprüchlich. Ich möchte sie deshalb nicht im einzelnen auseinanderlegen, sondern lediglich einige Punkte herausgreifen, die für meine Argumentation wichtig sind.

1) Wenn Mill (1861, 216) bei seiner Rede von Glück (happiness) zunächst die gleichen Wörter wie Bentham verwendet und Glück als Lust (pleasure) und Abwesenheit von Schmerz (pain) bestimmt, so gibt er dem Glücksbegriff doch im Verlauf seiner Darlegung eine ganz andere Bedeutung. „Glück“ ist bei Mill ein Terminus, der sich nicht unmittelbar auf Handlungsfolgen bezieht, sondern in der Beurteilung des Lebensablaufs eines Menschen zum Tragen kommt. Wenn wir auf ein gelungenes Leben zurückblicken können, wenn sich unsere Lebenspläne verwirklicht haben und wir auf Menschen getroffen

sind, denen wir Freude und Erfüllung verdanken und umgekehrt auch ihnen dies geben konnten, dann haben wir ein glückliches Leben gelebt. Wer nach Glück strebt, tut dies unter der Perspektive des gesamten Lebensverlaufs, er reflektiert in seiner Handlungsplanung immer zugleich die bisherige Lebensgeschichte und seine Lebensziele und sucht nicht nur die lustvolle Erregung eines Augenblicks (ebd., 215). In diesem Sinne kann Mill davon sprechen, daß Glück ein Endzweck ist¹⁶ und deshalb will er eine Handlung auch nicht danach beurteilen, ob sie mit ihren Folgen unmittelbar Lust oder Schmerz bringt, sondern ob sie eine Tendenz zeigt, das Glück zu befördern (ebd., 210).

2) Ähnlich wie Aristoteles betont Mill die aktive Seite eines glücklichen Lebens. Hierzu gehört die tätige Gestaltung des Lebens, durch die wir unsere Identität gewinnen und die gelungene Bewältigung der vielfältigen Herausforderungen und Aufgaben, die sich im privaten, beruflichen und öffentlichen Bereich stellen (S. 215, 217). Diesen Anforderungen unter Wahrung der eigenen Identität mit Würde gerecht zu werden, ist wesentlich für ein glückliches Leben. An dieser Stelle wird auch der Gegensatz von Glück und Zufriedenheit deutlich. Zufriedenheit ist eine statische Kategorie: wer zufrieden ist, hat seinen status quo mit denkbaren Alternativen verglichen und präferiert keine davon. Inhalt eines glücklicherfüllten Lebens kann jedoch gerade das Streben nach neuen Möglichkeiten und nach Vervollkommnung werden. Der Mill'sche Sokrates weiß sich auf dem richtigen Weg und ohne sein Ziel zu erreichen, kann er zugleich glücklich und doch unzufrieden sein (S. 212).

3) Noch in einem weiteren Punkt knüpft Mill direkt an Aristoteles an: Ein glückliches Leben erfordert die umfassende Nutzung all unserer Fähigkeiten und Möglichkeiten, je differenzierter wir unseren Verstand, unseren Geschmack und unser Empfindungsvermögen ausgebildet haben, umso reichhaltiger werden wir unser Leben gestalten können und umso vielfältigere Eindrücke werden wir aufnehmen können (S. 211).¹⁷ Von hier aus ist es auch verständlich, wenn Mill hinsichtlich der potentiellen Gegenstände unseres Glücksstrebens qualitative Unterschiede setzt und uns auffordert, nach dem zu streben, was uns umfassend beansprucht und eine differenzierte Weise der Aufnahme von uns verlangt. Insbesondere kann die Kultivierung unseres Gemeinschaftssinnes und unseres sozialen Mitgefühls zu einer entscheidenden Quelle des Glücks werden. In einer beeindruckenden Passage zeigt er, wie der Egoismus der Menschen die Freude am Leben ersticken kann: „When people who are tolerably fortunate in their outward lot do not find in life sufficient enjoyment to make it valuable to them, the cause generally is, caring for nobody but themselves. To those who have neither public nor private affections, the excitements of life are much curtailed, and in any case

dwindle in value as the time approaches when all selfish interests must be terminated by death: while those who leave after them objects of personal affection, and especially those who have also cultivated a fellow-feeling with the collective interests of mankind, retain as lively an interest in life on the eve of death as in the vigour of youth and health,“ (S. 215) Diese Stelle drückt zusammen mit einigen anderen (z.B. S. 218, 223) auf einen kurzen Nenner gebracht aus, daß der einzelne letztlich glücklich nur in einer glücklichen Gemeinschaft leben kann. Die kultivierte Anteilnahme am Schicksal der anderen macht es einem Menschen schwer, sein Glück in einer Welt von Leid und Not zu finden. Es ist nicht recht verständlich, warum Mill nicht dieses Argument zum Ausgangspunkt für seine Begründung der Forderung nach Maximierung des Gesamtwohls wählt und statt dessen einen gewundenen und letztlich irreführenden Weg zur Darlegung der Plausibilität beschreitet (ebd., Chap. IV). Es wäre sicher wesentlich überzeugender gewesen, zu argumentieren, daß das Streben nach Glück Kultivierung unseres Empfindungsvermögens voraussetzt, zu unseren Empfindungen auch das soziale Mitgefühl gehört und dessen Kultivierung eine rein selbstbezogene Haltung nicht zuläßt: das eigene Glück ist dann nicht von dem der anderen zu trennen und ein Streben nach möglichst viel Glück würde andere Menschen immer mit einbeziehen.

6. Einige Folgerungen

Wie wir gesehen haben, ist für Mill ein vernünftiges Streben nach Glück an eine Reihe von Voraussetzungen gebunden. Mit der gesellschaftlichen Bedingung der allgemeinen Rechtssicherheit haben wir bereits eine davon etwas näher kennengelernt. Daneben setzt er noch einige individuelle Bedingungen: Jeder Mensch muß über ein gewisses Maß an Bildung und Ausbildung verfügen, seine Sinne und Empfindungen müssen kultiviert sein und schließlich bedarf eine vernünftige Abwägung von Alternativen der Lebenserfahrung, durch die man die Folgen von Handlungen in ihren Auswirkungen auf andere Menschen erst richtig kennen und beurteilen lernt. (1861, 212; 1859, 262ff.)

So muß der Mensch vor einer verantwortlich zu nennenden Entscheidung ein umfassendes Erziehungsprogramm durchlaufen und zumindest, die Chance zur Aneignung von entsprechender Lebenserfahrung gehabt haben. Dies zusammen ermöglicht erst die Ausbildung einer individuellen Persönlichkeit und deshalb ist Mills Forderung (1861, 213) nur konsequent, daß niemand durch die Gesellschaft in eine Stellung oder zu einem Umgang gezwungen werden darf, worin bzw. wodurch seine Fähigkeiten verkümmern.

Anders als Bentham akzeptiert Mill die faktisch geäußerten Vorstellungen der Menschen vom glücklichen Leben nicht ohne weiteres, würde sich aber wiederum anders als Harsanyi keine Eingriffe in die Präferenzordnungen der Individuen erlauben. Hätte Mill das Wort „Präferenzautonomie“ verwendet, so hätte er es sicher nicht im Sinne einer undifferenzierten Gleichbehandlung faktischer Äußerungen gebraucht, sondern an die Erfüllung der Voraussetzungen für ein vernünftiges Entscheiden gebunden: Autonom handelt man nicht schon, wenn man verschiedene Alternativen in eine konsistente Reihenfolge bringen kann, sondern erst, wenn diese Reihenfolge auch begründet ist.

Dieses Verständnis von Autonomie hat unmittelbare Folgen für die Bestimmung der gesellschaftlichen Wohlfahrt: Wenn die utilitaristische Ethik nicht als eine institutionengebundene Ethik (Bentham, Neo-Utilitarismus) konzipiert, sondern mit Mill davon ausgegangen wird, daß die einzelnen Präferenzen unter Berücksichtigung der Interessen anderer und nicht willkürlich bestimmt werden, dann kann sich das Gesamtwohl nicht mehr einfach durch Aggregation von individuellen Präferenzordnungen (oder Nutzenfunktionen) ermitteln lassen. Das Aggregationsverfahren – sei es durch Abstimmung oder durch Summation – beruht in seinem Schematismus gerade in der Absehung von Gründen und macht somit die Bemühungen des einzelnen um diese Gründe sinnlos. Auf der anderen Seite kann man nicht davon ausgehen, daß jeder einzelne stets die Glücksvorstellungen der anderen richtig antizipieren wird und es deshalb zur Ermittlung des optimalen Gesamtwohls ausreichen würde, irgendeine individuelle Präferenzordnung herauszugreifen. Dies ginge nur, wenn die Menschen in identischer Weise denken und fühlen würden, ein Zustand, der nicht einmal normativ als anzustrebendes Ziel verstanden, mit dem Geist des Utilitarismus vereinbar wäre. Im Grunde genommen können nur die Vorstellungen derjenigen verlässlich antizipiert werden, mit denen man sich in einem engen Lebensverbund befindet. Eine größere Gemeinsamkeit kann nicht mehr antizipiert, sie muß erworben werden. Gerade unter diesem Aspekt bietet sich aber ein konsensethischer Ansatz, wie er z.B. in der konstruktiven Ethik vertreten wird, als eine kritische Weiterführung des Mill'schen Utilitarismus an.

Die Grundsituation, von der die konstruktive Ethik¹⁸ ausgeht, ist die eines Konflikts bei gleichzeitig bestehendem Interesse an der Aufrechterhaltung eines gemeinsamen Lebens- oder Handlungszusammenhangs. Von einem Konflikt sprechen wir, wenn von den Mitgliedern einer Gruppe Wünsche vorgebracht werden, die sich nicht gleichzeitig realisieren lassen, wobei der triviale Fall ausgeschlossen sein soll, daß sich ein Konflikt auflöst, indem eine Partei einen Wunsch durch einen anderen mit gleicher Präferenz ersetzt. Ein

„ernsthafter“ Konflikt läßt sich vernünftig, d.h. unter Wahrung des Interesses an Gemeinsamkeit lösen, wenn die einzelnen Konfliktparteien in einer Aussprache ihre Wünsche anmelden, ihre Gründe dafür vortragen und im Dialog bereit sind, ihre ursprünglichen Wünsche so zu modifizieren, daß ein von allen akzeptiertes Ergebnis zustande kommt. Ein solcher Dialog erfordert, daß jeder Teilnehmer die anderen als Person anerkennt, also bereit ist, ihre jeweiligen Gründe und die von ihnen vorgebrachte Kritik an der eigenen Position als Argumente ernstzunehmen. Das methodologische Prinzip, das einem solchen Dialog zugrunde liegt, läßt sich kurz so formulieren: „Transzendiere deine Subjektivität!“ Über dieses „Transsubjektivitätsprinzip“ lassen sich nun einige formale Forderungen für die Durchführung eines Dialogs begründen, deren Befolgung ihn dann als unverzerrt ausweisen. So muß er unvoreingenommen, zwanglos und nicht persuasiv geführt werden, d.h. niemand darf von der Teilnahme ausgeschlossen sein, Argumente und Personen dürfen nicht vorab positiv oder negativ ausgezeichnet sein und es dürfen keine Drohungen oder „strategische“ Argumente (Argumente wider besseres Wissen) vorgebracht werden. Wenn wir diesen Gedanken an die Überlegungen von Mill anschließen, dann können wir für das erste festhalten: Die Abgleichung unterschiedlicher Glücksvorstellungen muß im Rahmen eines unverzerrten Dialogs erfolgen, wobei die Bereitschaft, sich auf ihn einzulassen, nicht Ausdruck selbstbezogener Rationalität ist, sondern bereits in einer die anderen mit einschließenden Glücksvorstellung angelegt ist.

Auch für das Problem der Gerechtigkeit im engeren Sinn führt die Mill'sche Theorie des Glücks auf einen anderen als den von ihm selbst eingeschlagenen Weg. Die Menschen müssen zwar als gleich bezüglich ihrer Fähigkeit, nach dem utilitaristischen Programm zu handeln angesehen werden, jedoch als ungleich bezüglich der faktischen Ausbildung dieser Fähigkeit. Der Zugang zu Kultur und Bildung sowie zu den gesellschaftlichen Positionen, in denen sich Lebenserfahrung in einem anspruchsvollen Sinn erwerben läßt, steht in einer Gesellschaft nicht allen in gleicher Weise offen. Wer ein utilitaristisches Programm vertritt, muß also konsequent dafür eintreten, daß die Menschen in einen Stand versetzt werden, der es ihnen erlaubt, nach einem gelungenen Leben zu streben (wie dies Mill ja auch gefordert hat). Wenn wir aufgrund dieser Überlegungen als „Qualifikation“ im Sinne des „Hauptsatzes“ der Theorie der distributiven Gerechtigkeit die Ausbildung der Fähigkeit zu verantwortlicher Entscheidung betrachten, dann ist eine Gesellschaft gerecht im engeren Sinne, wenn sie ihren Mitgliedern gleiche Bildungs- und Ausbildungschancen¹⁹ gewährleistet sowie jedem die gleichen Möglichkeiten zur Wahrnehmung von Ämtern und Positionen einräumt. Diese Forderung selbst ist nicht abhängig von einer Bestätigung in einen unverzerrten Dialog,

da sie auf die Voraussetzungen abzielt, die den Dialog erst sinnvoll werden lassen. Je mehr ihr in einer Gesellschaft Rechnung getragen wird, desto umfassender läßt sich eine nicht bloß formale Abgleichung der Glücksvorstellungen erzielen. Prima facie sieht es nun so aus, als ob ein konsequent zu Ende gedachter Utilitarismus unmittelbar zum theoretischen Ansatz von Rawls führen würde. Denn was wir als Voraussetzungen für eine verantwortungsvolle Präferenzentscheidung angeführt haben, entspricht in etwa dem „greatest equal liberty principle“ und dem „principle of fair equality of opportunity“ (Rawls 1972, 60 u. 302).

Ich möchte den vielfältigen Beziehungen zwischen Utilitarismus und der Theorie der Gerechtigkeit von Rawls hier nicht mehr nachgehen²⁰ und so sei nur angedeutet, daß der erste Eindruck trügt. Der theoretische Kern des Rawls'schen Ansatzes, sein „difference principle“²¹, läßt sich nämlich in den utilitaristischen Rahmen, den Mill abgesteckt hat, nicht einpassen. Für Rawls ist wirtschaftliche und soziale Ungleichheit gerechtfertigt, wenn sie den am wenigsten Begünstigten den größtmöglichen Vorteil bringt.

So kann insbesondere aus einem Zustand der Gleichverteilung keine soziale oder wirtschaftliche Differenzierung mehr gerechtfertigt werden. Dies bedeutet aber eine ganz gravierende Einschränkung der Autonomie in ihrem soeben dargelegten Sinne. Selbst wenn man sich in einem unverzerrten Dialog, ausgehend von einer Gleichverteilungssituation, für eine ungleiche Verteilung aussprechen würde, müßte dieses Ergebnis als ungerecht im engeren Sinne abgelehnt werden. Rawls selbst kommt zur Aufstellung seines Differenzprinzips über die Frage nach Prinzipien, die freie und rationale, nur an ihrem eigenen Interesse orientierte Personen in einem ursprünglichen Zustand der Gleichheit wählen würden, um daran alle weiteren Vereinbarungen auszurichten. Wie wir jedoch gesehen haben, teilt Mill schon diese Ausgangsposition nicht, so daß die weiteren Diskrepanzen nicht allzu verwunderlich sind.²²

Natürlich ist nicht ausgeschlossen, daß das Differenzprinzip auch im Rahmen eines an Mill ausgerichteten utilitaristischen Programms als Regel politischen Handelns zur Geltung kommt, weil man z.B. durch seine Befolgung in kurzer Zeit zu einem Abbau von gesellschaftlichen Diskriminierungen kommen könnte. Aber es wäre dann in seiner Begründung an bestimmte historische Bedingungen gebunden und hätte somit gerade nicht den Status eines Prinzips, also eines apriorischen Satzes.

Ich hoffe, der Eindruck hat sich gefestigt, daß die Gerechtigkeitsproblematik im Utilitarismus noch lange nicht erschöpfend ausgelotet ist, wenngleich seine vulgären Varianten dies manchmal vermuten lassen. Auch mag die Erwartung etwas gedämpft worden sein, daß rein formale Methoden einen wesentlichen Beitrag zur Behandlung

dieses Problems leisten können. In jedem Falle sollte jedoch festgehalten werden, daß die Gerechtigkeit auch im Utilitarismus letzten Endes kein Mittel für die Herstellung eines Zustandes optimaler Wohlfahrt ist. Gerechtigkeit im engeren Sinne allein garantiert weder vernünftige Entscheidungen noch ein glückliches Leben, umgekehrt erhalten alle Bemühungen um eine gerechte Gesellschaftsordnung und alle Anstrengungen zu einem gerechten Handeln allein ihren Sinn durch die Vorstellung von einem glücklichen Leben in der Gemeinschaft, wobei dieses Streben zugleich selbst zur Quelle des Glücks werden kann. So hat also die Gerechtigkeit ihren Ort in der Topologie der praktischen Philosophie als die Bedingung der Möglichkeit zu einer menschenwürdigen Lebensführung.

Anmerkungen

- 1 Ich bin Herrn Prof. Dr. P. Lorenzen und Frau M. Erlwein-Kaiser für zahlreiche Anregungen und Hinweise dankbar.
- 2 K.J. Arrow (1963²). Arrow gab in der ersten Auflage seines Buches fünf Bedingungen an (vgl. 24, 26, 27, 29, 30), er hat sie in der zweiten Auflage modifiziert, vgl. 96. Wir beziehen uns auf die verbesserte Fassung.
- 3 K.J. Arrow (1963²), 59 und 96ff. Die Literatur zu und um diesen Satz ist kaum mehr zu übersehen. Elementare Einführungen sind Arrow (1969); Feldman (1980, Chap. 10); Mueller (1979, Chap. 10). Einen Überblick über die Entwicklung geben die Monographien von Sen (1970); Kelly, (1978); sowie die ausgezeichneten Übersichtsartikel von Sen (1977 und 1984).
- 4 Die grundlegenden Arbeiten hierzu sind Gibbard (1973) und Satterthwaite (1975). Elementare Einführungen in dieses Gebiet gibt Feldman (1979a, 1980). Einen Überblick zum Stand der Entwicklung findet man bei Pattanaik (1978) sowie in den genannten Übersichtsartikeln von Sen.
- 5 Eine Auswahlmenge $E(q)$ heißt pareto-optimal, wenn es zu keinem x aus $E(q)$ ein y aus A gibt mit $y R_i x$ für alle $p_i \in N$ und $y P_i x$ für wenigstens ein $p_i \in N$. Ein Beispiel: Sei $A = \{x, y, z\}$, $N = \{p_1, p_2\}$ und das Präferenzprofil q :

p_1	p_2
x	z
y	x
z	y

Dann ist die pareto-optimale Auswahlmenge $E(q) = \{x, z\}$. Für den Beweis, daß die Pareto-Regel strategiesicher und diktatorfrei ist vgl. Feldman (1979).

- 6 Interessant sind vor allem solche Versuche, die die Vorschläge von Rawls in die Social Choice Theory einbauen. Hier gibt es dann durchaus positive Resultate, vgl. z.B. Strasnick (1976 und 1979), Gibbard, (1979); zu einigen Problemen dieser Ansätze vgl. Kötter (1983).
- 7 Vgl. Neumann/Morgenstern (1947, 15-31, 617-632); ein ähnlicher Vorschlag stammt z.B. von Marschak (1950).
- 8 Der wesentliche Beitrag zu dieser Theorie stammt jedoch von H.R. Varian (1974). Eine gute Übersicht gibt E.A. Pazner (1976).
- 9 Varian (1974, 74) verwendet hier den Ausdruck "fair*", von „wealth-fair“ spricht er in Varian (1975).

- 10 Als weitere wichtige Beiträge zur Theorie der Fairness sind noch zu nennen: Feldman/Kirman (1974), Daniel (1975), Pazner (1977); einen Anwendungsversuch findet man bei Baumol (1982).
- 11 Wir haben bei Harsanyi einen Vorschlag in diese Richtung kennengelernt. Auch J.A. Mirrlees möchte Korrekturen an den Präferenzangaben einer Person zulassen, wenn sie dem „gesunden Menschenverstand“ entsprechen und sich Indizes dafür anführen lassen, daß die betreffende Person sich solchen Argumenten zugänglich erweist (Mirrlees 1982, 69).
- 12 Sen macht übrigens in einem Aufsatz nachdrücklich auf die negativen Konsequenzen eines zu allgemeinen Nutzenkonzepts aufmerksam (1977b, 335ff.).
- 13 Diese Formulierung wird Bentham zugeschrieben. Bei ihm selbst findet man folgenden Satz: „The happiness and unhappiness of any one member of the community – high or low, rich or poor – what greater or less part is it of the universal happiness and unhappiness, than that of any other?“ (1817, 459).
- 14 Dieser ihm zugeschriebene Ausspruch lautet genau: „Prejudice apart, the game of push-pin is of equal value with the arts and sciences of music and poetry.“ (1825, 253).
- 15 Es ist also ganz offensichtlich, daß Bentham eine an Institutionen gebundene Ethik vertritt: Als Privatmann mag jeder selbstbezogen handeln, in Erfüllung eines öffentlichen Amtes jedoch soll das Gebot der Wohlfahrtsmaximierung befolgt werden. Auch in diesem Punkt gibt es also eine enge Berührung mit dem Neo-Utilitarismus.
- 16 Diese Wendung führt Mill wiederholt an (vgl. 1861, 210, 213, 215; vgl. hierzu auch Brown (1973)).
- 17 Diesen Aspekt betont Mill auch in: (1859, Chap III, insbes. 262).
- 18 Zum Aufbau der konstruktiven Ethik vgl. z.B. Schwemmer, (1971, Kap. 4; Lorenzen, (1981, 49-63); Kambartel (1974). Eine gute Darstellung Konsens-ethischer Prinzipien (auch solcher von Habermas und Apel) findet sich bei Wimmer (1980, Erster Teil). Es sei übrigens ausdrücklich darauf hingewiesen, daß durch Konsens-ethische Prinzipien die Notwendigkeit von Abstimmungsverfahren nicht tangiert wird. Diskurse unterliegen i.d.R. Restriktionen, die eine zeitliche Beschränkung und damit u.U. den Abbruch der Diskurse erforderlich machen. Für diesen Fall müssen Verfahren zur Verfügung stehen, die dem Konsens-ethischen Prinzip „so weit wie möglich“ entsprechen, vgl. hierzu Kötter (1983). Interessant sind in diesem Zusammenhang auch die Bemerkungen, die Mill zur Moral öffentlicher Diskurse macht, vgl. Mill (1859, Chap. II).
- 19 „Ausbildung“ hat in diesem Zusammenhang nicht den engen Sinn von Berufsvorbereitung, vielmehr ist damit die Schulung aller Fähigkeiten gemeint, die der Mensch zur Lebensgestaltung gebrauchen kann.
- 20 Es sei hierzu ganz pauschal verwiesen auf die Reader von Daniels (1975), Höffe (1977); Blocker/Smith (1980).
- 21 Die „schlichte“ Fassung davon findet sich bei Rawls (1972, 60) (second principle, part (a)), die hier angesprochene auf 302.
- 22 In der Betonung der individuellen Rationalität und der Selbstbezogenheit stehen die Vertreter des Neo-Utilitarismus Rawls näher als Mill. Deshalb sind hier die Diskrepanzen auch nicht so fundamental wie im Verhältnis von Rawls zu Mill; vgl. hierzu z.B. Yaari (1981); Gaertner (1981).

Literaturverzeichnis

- Arrow, K.J. (1951), (1963²), *Social Choice and Individual Values*, New Haven & London.
- (1969), *Values and Collective Decision-Making*, in: P. Laslett/W.G. Runciman (Eds.), *Philosophy, Politics and Society*, Oxford 1969, 215-232
- Baumol, W.J. (1982), *Applied Fairness Theory and Rationing Policy*, AER 72, 639-651
- Bedau, H. (1963), *Justice and Classical Utilitarianism*, Nomos 6, 284-305
- Bentham, J., (1817), *Plan of Parliamentary Reform, in the Form of a Catechism, with Reasons for each Article: with an Introduction showing the Necessity of Radical, and the Inadequacy of Moderate Reform*, London (in: J. Bentham, Works (Ed. J. Bowring), vol. III, Edinburgh 1843)
- (1823), *An Introduction to the Principles of Morals and Legislation*, London (Repr. N.Y. 1948)
- (1825), *Rationale of Reward*, London (in: J. Bentham, Works (Ed. J. Bowring), vol. II, Edinburgh 1843)
- Blocker, G./Smith, E.H. (Eds.), (1980), *John Rawls Theory of Social Justice*, Athens
- Brown, D.G., (1973), *What is Mill's Principle of Utility?*, Can. J. Phil. 3, 1-12
- Daniel, T.E., (1977), *A Revised Concept of Distributional Equity*, J. Econ. Th. 14, 458-466
- Daniels, N. (Ed.), (1975), *Reading Rawls. Critical Studies of 'A Theory of Justice'*, Oxford
- Debreu, G. (1959), *Theory of Value*, New York
- Diamond, P. (1967), *Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility: A Comment*, J.Pol.Ec. 75, 765-766
- Feldman, A.M./Kirman, A., (1974), *Fairness and Envy*, AER 64, 995-1005
- Feldman, A.M., (1979a), *Manipulating Voting Procedures*, Econ. Inquiry 17, 452-474
- (1979b), *Manipulation and the Pareto Rule*, J. Econ. Th. 21, 473-482
- (1980), *Welfare Economics and Social Choice Theory*, Boston u.a.
- Foley, D. (1967), *Resource Allocation in the Public Sector*, Yale Econ. Essays 7, 45-98
- Gärdenfors, P. (1978), *Fairness without interpersonal Comparisons*, Theoria 44, 57-74
- (1979), *On Definitions of Manipulation of Social Choice Functions*, in: J.-J. Laffont (Ed.), *Aggregation and Revelation of Preferences*, Amsterdam u.a., 29-36
- Gaertner, W. (1981), *Rawlsianism, Utilitarianism, and Profiles of Extended Orderings*, Z. ges. Staatswiss. 137, 78-96
- Gaus, G. (1980), *Mill's Theory of Moral Rules*, Australasian J. Phil. 58, 265-279
- Gibbard, A. (1973), *Manipulation of Voting Schemas: A General Result*, Econometrica 41, 587-601
- (1978), *Disparate Goods and Rawls' Difference Principle: A Social Choice Theoretic Treatment*, Theory and Decision 11, 267-288
- Harsanyi, J. (1976), *Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility*, in: ders., *Essays on Ethics, Social Behavior, and Scientific Explanations*, Dordrecht & Boston, 6-23 (ursprüngl. in J.Pol.Ec. 63 (1955), 309-321)
- (1977a), *Bayesian Decision Theory and Utilitarian Ethics*, AER (Pap. & Proc.) 68, 223-228
- (1977b), *Morality and the Theory of Rational Behavior*, Social Research 44, 623-656

- Höffe, O. (Hrsg.), (1977), *Über John Rawls' Theorie der Gerechtigkeit*, Frankfurt
 – (1979), *Zur Theorie des Glücks im klassischen Utilitarismus*, in: ders., *Ethik und Politik*, Frankfurt, 120-159
- Kambartel, F. (1974), *Wie ist praktische Philosophie konstruktiv möglich?* in: ders., *Praktische Philosophie und Konstruktive Wissenschaftstheorie*, Frankfurt, 9-33
- Kelly, J.S., (1978), *Arrow Impossibility Theorems*, New York
- Kötter, R., (1983), *Effizienz, Rationalität und Verteilungsgerechtigkeit – Zum normativen Hintergrund der Betriebswirtschaftslehre*, in: E. Kappler (Hrsg.), *Rekonstruktion der Betriebswirtschaftslehre als ökonomische Theorie*, Spardorf, 277-302
- Lorenzen, P. (1981), *Politische Anthropologie*, in: O. Schwemmer (Hrsg.), *Vernunft, Handlung und Erfahrung*, München, 49-63
- Lyons, D. (1978), *Mill's Theory of Justice*, in: A.J. Goldman/J. Kim (Eds.), *1-20, Values and Morals*, Dordrecht & Boston, 1-20
- Marschak, J. (1950), *Rational Behavior, Uncertain Prospects and Measurable Utility*, *Econometrica* 18, 111-114
- McClennen, E.F. (1981), *Constitutional Choice: Rawls versus Harsanyi*, in: J. Pitt (Ed.), *Philosophy in Economics*, Dordrecht & Boston, 93-110
- Mill, J. (1833), *Fragment on Mackintosh*, London
- Mill, J.St. (1859), *On Liberty*, London (in: J.St. Mill, *Coll. Works* (Eds. F.E.L. Priestley/J.M. Robson e.a.) vol. XIIX, Toronto 1977)
- (1861), *Utilitarianism*, London (in: J.St. Mill, *Coll. Works* (Eds. F.E.L. Priestley/J.M. Robson), vol. X, Toronto 1969)
- Mirrlees, J.A. (1982), *The Economic Uses of Utilitarianism*, in: A. Sen/B. Williams (Eds.), *Utilitarianism and Beyond*, Cambridge, 63-84
- Mueller, D.C. (1979), *Public Choice*, Cambridge
- Neumann, J.v./O. Morgenstern (1944), (1947²), *Theory of Games and Economic Behavior*, Princeton
- Pattanaik, P.K., (1978), *Strategy and Group Choice*, Amsterdam u.a.
- Pazner, E.A., (1976), *Recent Thinking on Economic Justice*, *J. Peace Sci.* 2, 143-154
- /Schmeidler, D. (1978), *Egalitarian Equivalent Allocations: A New Concept of Economic Equity*, *Quart.J.Ec.* 93, 671-687
- Rawls, J., (1972), *A Theory of Justice*, Harvard (Oxford 1972)
- Satterthwaite, M.A., (1975), *Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions*, *J.Econ.Th.* 10, 187-217
- Schwemmer, O. (1971), (1980²), *Philosophie der Praxis*, Frankfurt
- Sen, A. (1970), *Collective Choice and Social Welfare*, San Francisco & London
- (1977a), *Social Choice Theory: A Re-examination*, *Econometrica* 45, 53-89
- (1977b), *Rational Fools: a Critique of the Behavioral Foundations of Economic Theory*, *Phil. & Public Affairs* 6, 317-344
- (1979), *Utilitarianism and Welfareism*, *J.Phil.* 76, 463-489
- (1984), *Art. „Social Choice Theory“* in: K.J. Arrow/M. Intriligator (Eds.), *Handbook of Mathematical Economics*, vol. III, Amsterdam u.a.
- Strasnick, St. (1976), *The Problem of Social Choice: Arrow to Rawls*, *Philosophy and Public Affairs* 5, 241-273
- (1979), *Moral Structures and Axiomatic Theory*, *Theory and Decision* 11, 195-206
- (1981), *Neo-Utilitarian Ethics and the Ordinal Representation Assumption*, in: J. Pitt (Ed.), *Philosophy in Economics*, Dordrecht & Boston, 63-92
- Varian, H.R. (1974), *Equity, Envy, and Efficiency*, *J. Econ. Th* 9, 63-91

- (1975), *Distributive Justice, Welfare Economics, and the Theory of Fairness*, *Phil. Public Affairs* 4, 223-247
- Wimmer, R., (1980), *Universalisierung in der Ethik*, Frankfurt
- Yaari, M.E., (1981), *Rawls, Edgeworth, Shapley, Nash: Theories of Distributive Justice Re-examined*, *J.Econ.Th.* 24, 1-39.