

## Texte zum Utilitarismus



# Texte zum Utilitarismus

Herausgegeben von Jörg Schroth

Reclam

RECLAMS UNIVERSAL-BIBLIOTHEK Nr. 19350  
2016 Philipp Reclam jun. GmbH & Co. KG, Stuttgart  
Gestaltung: Cornelia Feyll, Friedrich Forssman  
Gesamtherstellung: Reclam, Ditzingen. Printed in Germany 2017  
RECLAM, UNIVERSAL-BIBLIOTHEK und  
RECLAMS UNIVERSAL-BIBLIOTHEK sind eingetragene Marken  
der Philipp Reclam jun. GmbH & Co. KG, Stuttgart  
ISBN 978-3-15-019350-1  
[www.reclam.de](http://www.reclam.de)



# Inhalt

Einleitung 7

JEREMY BENTHAM

Aus: Eine Einführung in die Prinzipien der Moral und  
der Gesetzgebung 32

JOHN STUART MILL

Aus: Der Utilitarismus 51

HENRY SIDGWICK

Aus: Die Methoden der Ethik 92

G. E. MOORE

Aus: Principia Ethica 141

G. E. MOORE

Aus: Grundprobleme der Ethik 169

R. M. HARE

Ethische Theorie und Utilitarismus 178

PETER SINGER

Aus: Praktische Ethik 208

DIETER BIRNBACHER

Eine Begründung des Utilitarismus 211

JOHN RAWLS

Aus: Eine Theorie der Gerechtigkeit 238

WILL KYMLICKA

Aus: Politische Philosophie heute 245

BERNARD WILLIAMS

Aus: Kritik des Utilitarismus 254

Anhang

Literaturhinweise 301

Textnachweise 331

# Einleitung

Die wichtigen Fragen, die der Utilitarismus stellt, sollten in Zusammenhängen erörtert werden, die vielversprechender sind als die des Utilitarismus. Der Tag kann nicht mehr allzuweit entfernt sein, an dem wir nichts mehr von ihm hören.

Bernard Williams (1973)<sup>1</sup>

Späteren Generationen wird ein Großteil der Moralphilosophie des zwanzigsten Jahrhunderts wie ein Kampf erscheinen, um dem Utilitarismus zu *entkommen*. Kaum scheint es uns gelungen, eine utilitaristische Lehre zu widerlegen, finden wir uns schon gefangen im Griff der nächsten wieder.

Christine Korsgaard (1993)<sup>1</sup>

Von diesen Vorhersagen zweier Gegner des Utilitarismus hat sich Williams' Vorhersage als Fehleinschätzung erwiesen, während Korsgaards Diagnose die Moralphilosophie des 20. Jahrhunderts treffend charakterisiert. Allen Einwänden und Widerlegungsversuchen zum Trotz ist der Utilitarismus aus der Ethik nicht mehr wegzudenken. Man kann sogar den Eindruck gewinnen, dass der Utilitarismus die *default theory* ist und jede Abweichung von ihm einer besonderen Rechtfertigung bedarf – eine Auffassung, die in diesem Band explizit von Peter Singer vertreten wird.<sup>2</sup>

1 Bernard Williams, *Kritik des Utilitarismus*, in diesem Band, S. 298, und Christine M. Korsgaard, *The Reasons We Can Share*, in: Christine M. Korsgaard, *Creating the Kingdom of Ends*, Cambridge 1996, S. 275.

2 Dabei wird hier die zweite Auflage seiner *Praktischen Ethik* zitiert (obwohl bereits die dritte Auflage vollständig überarbeitet erschienen ist), weil Singer dort sein Argument für den Utilitarismus deutlicher formuliert.

Im Laufe seiner Geschichte hat sich der Utilitarismus in unterschiedliche Varianten ausdifferenziert, die ein so weites Spektrum umfassen, dass es kaum möglich ist, sie in einer auf alle Varianten zutreffenden Definition zusammenzufassen. Eine Idee jedoch ist den meisten Varianten des Utilitarismus gemeinsam: In der Moral geht es nur um eines, nämlich um die Förderung des Guten, genauer: um die Förderung des Wohlbefindens der Menschen (bzw. aller empfindungsfähigen Wesen). Zwar ist selbst diese Idee nicht die Grundidee aller Utilitarismen, also eine Idee, die alle Varianten motiviert und von der alle Begründungen ausgehen – in R. M. Hares und Peter Singers Begründung des Utilitarismus (in diesem Band) spielt sie keine Rolle. Sie ist aber zumindest, als Ausgangsidee oder als Folgerung aus anderen Ideen, den meisten Varianten des Utilitarismus gemeinsam. Mit dieser Idee ist der Utilitarismus nicht hinreichend genau und erschöpfend charakterisiert. Für ein genaueres Bild müssen wir uns die wesentlichen Behauptungen und Bestandteile des Utilitarismus vor Augen führen.

### 1. Theorie des Guten und Welfarismus

Da es für Utilitaristen in der Moral nur um die Förderung des Guten geht, ist die **Theorie des Guten** (auch »**Werttheorie**« oder »**Axiologie**« genannt) ein zentraler Bestandteil des Utilitarismus. Sie bestimmt, was das **intrinsisch Gute**, also das um seiner selbst willen (und nicht bloß als Mittel für andere Zwecke) Schätzens- und Erstrebenswertes ist. Statt vom Guten sprechen Utilitaristen auch vom Nutzen (*utility*), Glück oder Wohlbefinden (*well-being*, *welfare*). Diese Begriffe sind weniger allgemein als der Begriff des Guten, da sie das Gute auf das Gute für empfindungsfähige Wesen einschränken, aber immer noch wenig aussagekräftig. In ihrer Erklärung, was un-

ter Glück und Wohlbefinden zu verstehen ist bzw. was unser Glück und Wohlbefinden ausmacht, unterscheiden sich die verschiedenen von Utilitaristen vertretenen Theorien des Guten.<sup>3</sup>

Die drei klassischen Utilitaristen, Jeremy Bentham (1748–1832), John Stuart Mill (1806–1873) und Henry Sidgwick (1838–1900), vertraten in ihrer Theorie des Guten einen **Hedonismus**, dem zufolge Freude (*pleasure*) und die Vermeidung von Schmerzen bzw. Leid das einzige um seiner selbst willen Erstrebenswerte ist.

Zahlreiche zeitgenössische Utilitaristen – in diesem Band R. M. Hare (1919–2002) und Peter Singer (\*1946) – haben den Hedonismus zugunsten einer **Wunsch- bzw. Präferenztheorie des Guten** aufgegeben. Für sie besteht das Gute nicht in Freude, sondern in der Erfüllung unserer Wünsche bzw. Präferenzen: Je weitergehend unsere Präferenzen erfüllt werden, desto besser verläuft unser Leben.

Andere Utilitaristen sind der Meinung, dass es bestimmte objektive Güter gibt, die ein notwendiger Bestandteil eines guten Lebens sind. Diese Güter werden in einer Liste aufgezählt. Man spricht deshalb von einer **objektiven Listentheorie des Guten** (*objective list theory*). Beispielsweise steht auf James Griffins Liste:<sup>4</sup> Vergnügen (*enjoyment*), etwas mit seinem Leben erreichen (*accomplishment*), tiefe persönliche Beziehungen, Verständnis (*understanding*) über uns selbst und unseren Platz in der Welt sowie die Elemente einer charakteristisch menschlichen Existenz: Autonomie, die zum Handeln

3 Vgl. zu den folgenden Theorien des Guten: Chris Heathwood, Subjective theories of Well-Being, in: *The Cambridge Companion to Utilitarianism*, hrsg. von Ben Eggleston und Dale E. Miller, Cambridge 2014, S. 199–219; Ben Bradley, Objective theories of well-being, ebd., S. 220–238.

4 James Griffin, *Value Judgment*, Oxford 1996, S. 29 f.

notwendigen Fähigkeiten sowie Freiheit. Auch G. E. Moore (1873–1958) kritisierte den Hedonismus heftig und vertrat eine aus zwei Elementen bestehende Listentheorie:

Die bei weitem wertvollsten Dinge, die wir kennen oder uns vorstellen können, sind gewisse Bewußtseinszustände, die sich summarisch umschreiben lassen als die Freuden [*pleasures*] menschlichen Umgangs und das Genießen schöner Gegenstände. Wahrscheinlich hat niemand, der sich die Frage vorgelegt hat, je daran gezweifelt, daß persönliche Zuneigung und die Wertschätzung des Schönen in Kunst oder Natur gut an sich sind; noch dürfte es wahrscheinlich sein, daß jemand bei genauer Überlegung, welche Dinge *rein um ihrer selbst willen* den Besitz wert sind, glauben kann, irgend etwas anderes habe *annähernd* soviel Wert wie die unter diesen zwei Bezeichnungen zusammengefaßten Dinge.<sup>5</sup>

Moores Utilitarismus wurde als **idealer Utilitarismus** bezeichnet. Heute bezeichnet man einen Utilitarismus mit einer objektiven Listentheorie des Guten auch als **objektiven Utilitarismus** (im Unterschied zum **subjektiven Utilitarismus** mit einer Präferenztheorie oder hedonistischen Theorie des Guten) oder, englisch, *objective list utilitarianism*. Da der Begriff »objektiver Utilitarismus« noch in anderen Bedeutungen verwendet wird, sollte man ihn jedoch besser als **Güterlistenutilitarismus** bezeichnen.

Die drei Theorien des Guten sind selbst keine utilitaristischen Theorien, d. h. Theorien, die utilitaristische Annahmen voraussetzen oder nur von Utilitaristen vertreten werden können. Auch sind sie, obwohl Bestandteil moralischer Theorien,

5 G. E. Moore (1903), *Principia Ethica*, erweiterte Ausgabe, Stuttgart 1996, S. 260 (§ 113).

selbst keine *moralischen* Theorien: Hedonismus, Präferenztheorie und objektive Listentheorie bestimmen, *ohne dabei auf moralische Erwägungen zurückzugreifen*, was um seiner selbst willen für Menschen schätzens- und erstrebenswert ist: Es handelt sich um Theorien über das *gute Leben*, nicht aber um Theorien über das *moralisch gute Leben*. Man spricht daher auch von Theorien des nicht-moralisch bzw. außermoralisch Guten, und es ist charakteristisch für den Utilitarismus, dass sein Ziel nicht die Förderung moralischer Güter und Werte wie z. B. Gerechtigkeit, Ehrlichkeit etc. ist, sondern nur die **Förderung des außermoralisch Guten**. Diese Beschränkung auf das außermoralisch Gute kennzeichnet den Utilitarismus als eine **teleologische Theorie** (vgl. S. 239 ff.).

Oft stößt man im Zusammenhang mit dem Utilitarismus auf den Begriff »**Welfarismus**« (*welfarism*). Der Welfarismus ist in der Tat charakteristisch für den Utilitarismus.<sup>6</sup> Allerdings wird dieser Begriff in zwei unterschiedlichen, nicht immer explizit unterschiedenen Bedeutungen verwendet:

1. Welfarismus bezeichnet die Auffassung, dass es in der Moral nur um die Förderung des Wohlbefindens geht, dass also Wohlbefinden das einzige *moralisch relevante* Gute ist. Mit dieser Auffassung ist vereinbar, dass es noch andere intrinsische (also an sich gute) Güter gibt, z. B. das Schöne bzw. ästhetische Werte, die intrinsisch gut sind, unabhängig davon, ob sie zu unserem Wohlbefinden beitragen oder nicht. Der Welfarismus in dieser Bedeutung bestreitet jedoch, dass es

6 Vgl. zum Welfarismus: Simon Keller, *Welfarism*, *Philosophy Compass* 4 (2009), S. 82–95; Nils Holtug, *Welfarism – The Very Idea*, *Utilitas* 15 (2003), S. 151–174; Andrew Moore / Roger Crisp, *Welfarism in Moral Theory*, *Australasian Journal of Philosophy* 74 (1996), S. 598–613.

in der Moral um die Förderung solcher Güter geht. So verstanden, entspricht der Welfarismus der eingangs genannten gemeinsamen Idee fast aller Varianten des Utilitarismus.

2. Welfarismus bezeichnet die Auffassung, dass Wohlbefinden das einzige intrinsisch Gute ist. Das Schöne wäre dieser Auffassung zufolge kein eigenständiges intrinsisches Gut, sondern nur wertvoll, insofern es zu unserem Wohlbefinden beiträgt oder ein Mittel dazu ist. Obwohl Utilitaristen den Welfarismus auch in dieser Bedeutung vertreten, muss man dies nicht tun, um Utilitarist zu sein: Man könnte die Auffassung vertreten, dass es zwar neben Wohlbefinden noch andere intrinsische Güter gibt, diese aber in der Moral keine (bzw. nur eine indirekte) Rolle spielen.

Entsprechend der ersten Bedeutung ist der Welfarismus keine Theorie des Guten, sondern eine Auffassung darüber, worum es in der Moral geht, also eine Auffassung über Sinn und Zweck der Moral. In der zweiten Bedeutung ist er eine Theorie des Guten, und die drei spezielleren Theorien des Guten (Hedonismus, Präferenztheorie und objektive Listentheorie) sind welfaristische Theorien des Guten.

## 2. Aggregation und bester Zustand: Aufsummierung des Wohlbefindens einzelner Individuen zum Gesamtnutzen

Die erwähnten Theorien des Guten bestimmen, was ein einzelnes menschliches Leben zu einem guten Leben macht. Es handelt sich also um Theorien über das persönliche, individuelle Gute bzw. über das Gute für einzelne Individuen. In der Moral geht es jedoch – wenn wir offenlassen, ob es auch Pflichten gegenüber sich selbst gibt – um den Umgang mit anderen Menschen, also um das moralisch richtige Verhalten in der Interaktion von mindestens zwei Menschen. Es geht um mora-

lisch richtiges Verhalten, wenn das Wohlbefinden mehrerer Menschen betroffen ist. Wie bestimmt man, was das Beste bzw. der beste Zustand für eine Gruppe von Menschen ist? Bei einem einzelnen Menschen ist offensichtlich, was das Beste für ihn ist. Getreu dem Prinzip »Je mehr von einem Gut, desto besser!« – das als Rationalitätsprinzip oder als ein aus dem Begriff des Guten folgendes Prinzip verstanden werden kann – gilt: Je mehr Freude man im Leben empfindet, je mehr Präferenzen erfüllt werden oder je mehr Güter der Güterliste realisiert werden (und je höher das Ausmaß der Realisierung der einzelnen Güter ist), desto besser.

Utilitaristen nehmen an, dass es vernünftig ist, dieses Prinzip auf eine Mehrzahl von Menschen zu übertragen. Sie nehmen an, dass man das Wohlbefinden verschiedener Menschen miteinander vergleichen und zu einer Gesamtmenge an Wohlbefinden aufsummieren kann und dass **der beste Zustand** derjenige ist, in dem der **Gesamtnutzen**, d. h. die Gesamtsumme an Wohlbefinden abzüglich der Gesamtsumme an Leid, am größten ist. Diese Auffassung vom besten Zustand ergibt den **Nutzensummenutilitarismus** (*total utilitarianism*) – die Standardform des Utilitarismus.

Eine Folge dieser Auffassung ist, dass es keine Rolle spielt, ob der größte Gesamtnutzen gleich oder ungleich auf einzelne Individuen verteilt ist. Auch kann der beste Zustand ein Zustand sein, in dem es einigen Menschen (darunter dem Akteur, falls er für die Herbeiführung des besten Zustands ein großes Opfer erbringen musste) sehr schlecht geht. Außerdem impliziert der Nutzensummenutilitarismus die von Parfit so genannte »*repugnant conclusion*«<sup>7</sup>: Eine Welt mit einer riesigen Bevölkerungszahl, in der es den Menschen so schlecht geht, dass ihr Leben gerade noch lebenswert ist, ist besser als eine

7 Derek Parfit, *Reasons and Persons*, Oxford 1984, S. 388.

Welt mit weniger Menschen, denen es allen sehr gut geht, sofern die Nutzensumme in der bevölkerungsreichen Welt größer ist als in der anderen Welt. Ferner folgt aus dieser Auffassung über den besten Zustand zusammen mit dem Gebot, den besten Zustand herbeizuführen, eine Fortpflanzungspflicht.

Solche kontraintuitiven Implikationen könnte man mit einem **Durchschnittsnutzenutilitarismus** (*average utilitarianism*) vermeiden, dem zufolge der Zustand mit dem größten Durchschnittsnutzen der beste Zustand ist. Doch auch dieser hat kontraintuitive Implikationen, z. B. dass es richtig ist, wenn eine Bevölkerungsgruppe, deren Wohlbefinden unter dem Durchschnitt liegt, ausgelöscht wird.

Ein Ausweg aus diesen Schwierigkeiten ist der **person-affecting utilitarianism**, dem zufolge es in der Moral nur um bereits existierende (bzw. entstehende) Menschen geht und es nicht unsere Aufgabe ist, glückliche Menschen zu machen, sondern Menschen glücklich zu machen. Ungelöst bleibt dabei das Problem, wie man im Rahmen des Utilitarismus bevölkerungspolitische Fragen und Fragen nach dem Umgang mit zukünftigen Generationen beantworten soll.<sup>8</sup>

8 Vgl. zur *repugnant conclusion*, zur Unterscheidung zwischen Nutzensummen- und Durchschnittsnutzenutilitarismus sowie zum *person-affecting utilitarianism*: Tim Mulgan, *Utilitarianism and Our Obligations to Future People*, in: *The Cambridge Companion to Utilitarianism*, hrsg. von Ben Eggleston und Dale E. Miller, Cambridge 2014, S. 325–347; Gustaf Arrhenius / Jesper Ryberg / Torbjörn Tännsjö, *The Repugnant Conclusion*, in: *Stanford Encyclopedia of Philosophy*, hrsg. von Edward N. Zalta, <http://plato.stanford.edu/entries/repugnant-conclusion> (2010); Krister Bykvist, *Utilitarianism. A Guide for the Perplexed*, London 2010, S. 62–67; Tim Mulgan, *Understanding Utilitarianism*, Stocksfield 2007, S. 170–175, Tim Mulgan, *Future People. A Moderate Consequentialist Account of Our Ob-*

### 3. Universalismus

Im Utilitarismus zählt *das Gute aller Wesen* bei der Berechnung der Nutzensumme. Der Utilitarismus dehnt den Schutzbereich der Moral auf alle Wesen aus, denen man ein Wohlbefinden im Sinne der erwähnten Theorien des Guten zuschreiben kann. Alle empfindungsfähigen Tiere fallen somit in den Schutzbereich der Moral, so dass bei der Frage, welcher Zustand (in einer bestimmten Situation) der beste ist, auch das Wohlbefinden der betroffenen Tiere berücksichtigt werden muss.

### 4. Unparteilichkeit

Im Utilitarismus zählt das Gute aller Wesen *gleich viel* bei der Berechnung der Nutzensumme. Es dürfen daher z. B. niemandes Präferenzen bei der Berechnung der Nutzensumme unberücksichtigt bleiben oder mehr oder weniger zählen als die Präferenzen anderer Menschen. Insbesondere zählen meine eigenen Präferenzen und die Präferenzen mir nahestehender Menschen nicht mehr als die Präferenzen anderer Menschen. Ein berühmter Slogan für diese Unparteilichkeit ist »Benthams Diktum«: »Jeder zählt für einen, keiner für mehr als einen.«<sup>9</sup>

### 5. Akteur-Neutralität

Im Utilitarismus geht es nicht darum, dass *ich* bestimmte Dinge tue oder nicht tue, sondern darum, dass bestimmte Dinge geschehen oder nicht geschehen, je nachdem, ob sie die Nut-

*ligations to Future Generations*, Oxford 2006; *The Repugnant Conclusion. Essays on Population Ethics*, hrsg. von Jesper Ryberg und Torbjörn Tännsjö, Dordrecht 2004.

9 John Stuart Mill (1861), *Der Utilitarismus*, Stuttgart 1976, S. 108; zweisprachige Neuausgabe: Stuttgart 2006, S. 185.

zensumme erhöhen oder mindern. Alle utilitaristischen Akteure haben daher das gleiche Ziel, nämlich für die Maximierung der Nutzensumme zu sorgen. Da z. B. Töten und das Brechen von Versprechen meistens schlechte Konsequenzen haben und die Nutzensumme mindern werden, ist es Aufgabe jedes Akteurs, dafür zu sorgen, dass möglichst wenige dieser Handlungen geschehen. Utilitaristische Gebote sind daher **akteur-neutrale Gebote**: Ich soll dafür sorgen, dass möglichst wenige Versprechen gebrochen werden und möglichst wenige Menschen getötet werden.

Meistens kann ich diese Gebote befolgen, indem ich selbst meine Versprechen nicht breche und selbst nicht töte. Gerade ich jedoch in eine Situation, in der nur dadurch, dass ich einen Menschen töte, verhindert werden kann, dass mehrere Menschen getötet werden, gebietet mir das akteur-neutrale Gebot, zu töten.<sup>10</sup>

Moralische Gebote und Verbote der herkömmlichen Moral sind dagegen **akteur-relative Gebote** und Verbote: Ich soll *meine* Versprechen halten, ich soll für *meine* Kinder sorgen, ich darf *selbst* nicht töten. Akteur-relative Verbote (im Englischen auch »*constraints*« genannt) verbieten (absolut oder bis zu einem Schwellenwert), dass ich ein Versprechen breche, um die Zahl der Versprechensbrüche zu minimieren, oder dass ich jemanden töte, um die Zahl der Tötungen zu minimieren.

10 Vgl. zur Unterscheidung zwischen Akteur-Relativität und Akteur-Neutralität: Michael Ridge, Reasons for Action: Agent-Neutral vs. Agent-Relative, in: *Stanford Encyclopedia of Philosophy*, hrsg. von Edward N. Zalta, <http://plato.stanford.edu/entries/reasons-agent> (2011).

## 6. Konsequentialismus

Der Utilitarismus ist eine **konsequentialistische Theorie**, also eine Theorie, in der die Richtigkeit bzw. Falschheit einer Handlung nur von den Konsequenzen der Handlung abhängt: Da die Auswirkung auf das Wohlbefinden aller Betroffenen eine Konsequenz einer Handlung und nichts anderes moralisch relevant ist, hängt die Richtigkeit im Utilitarismus nur von den Konsequenzen ab – und zwar auch von den in fernster Zukunft liegenden Konsequenzen.

Dabei sind aufgrund des Welfarismus nur diejenigen Konsequenzen moralisch relevant, die Auswirkungen auf das Wohlbefinden haben. Verletzt man z. B. die Würde einer Person, ist die Würdeverletzung zwar eine Konsequenz einer Handlung, aber diese Konsequenz ist für sich genommen nicht moralisch relevant. Moralisch relevant ist nicht die Würdeverletzung als solche, also *dass* die Würde einer Person verletzt wurde, sondern nur, welche Auswirkungen die Würdeverletzung auf das Wohlbefinden der betroffenen Menschen hat.

Uneins sind sich Utilitaristen in der Frage, ob die Richtigkeit einer Handlung von den tatsächlichen oder von den wahrscheinlich zu erwartenden Konsequenzen abhängt. Entsprechend unterscheidet man zwischen einem ***actual consequence utilitarianism*** bzw. **objektiven Utilitarismus** und einem ***probable consequence utilitarianism*** bzw. **subjektiven Utilitarismus**.

## 7. Maximierung

Die meisten Utilitaristen vertreten die Auffassung, dass man das Gute *maximieren* soll: Moralisch richtig ist also diejenige Handlung, die den Zustand mit dem *größten* Gesamtnutzen zur Folge hat. Allgemeiner formuliert: Moralisch richtig ist die

Handlung mit den *besten* Konsequenzen, also die Handlung, die den *besten* Zustand zur Folge hat. Dies ergibt folgendes **Richtigkeitskriterium**:

Eine Handlung ist moralisch richtig genau dann, wenn es keine andere Handlung gibt (die die handelnde Person ausführen könnte), die einen größeren Gesamtnutzen zur Folge hat.

Da die Richtigkeit von den Konsequenzen von *Handlungen* abhängt, bezeichnet man einen Utilitarismus mit diesem Richtigkeitskriterium als **(maximierenden) Handlungsutilitarismus** (*maximizing act-utilitarianism*). Er stellt die von den meisten Utilitaristen vertretene Theorie dar, und spricht man vom Utilitarismus, meint man in der Regel den maximierenden Handlungsutilitarismus.

8. Richtigkeitskriterium vs. Entscheidungskriterium:  
Direkter vs. indirekter Handlungsutilitarismus

Formuliert man das Richtigkeitskriterium als Gebot, erhält man folgendes **Entscheidungskriterium**:

Führe in jeder Situation nur diejenige Handlung aus, zu der es keine andere Handlungsalternative mit größerem zu erwartenden Gesamtnutzen gibt.

Einen Handlungsutilitarismus mit diesem Entscheidungskriterium bezeichnet man als **direkten Handlungsutilitarismus**, da das Richtigkeitskriterium in jeder Situation *direkt* als Entscheidungskriterium angewendet wird.

Gegen dieses Entscheidungskriterium sprechen eine Reihe von Einwänden, die darauf hinauslaufen, dass es nicht zum

größten Gesamtnutzen führt, wenn alle stets nach diesem Entscheidungskriterium handeln.<sup>11</sup> Denn um es anzuwenden, müssten wir alle uns offenstehenden Handlungsalternativen herausfinden, zu jeder dieser Handlungsalternativen alle möglicherweise eintretenden Konsequenzen ermitteln, jede dieser möglichen Konsequenzen bewerten und jeder dieser möglichen Konsequenzen eine Eintrittswahrscheinlichkeit zuordnen. Damit wären wir heillos überfordert: Wir sind (aus Zeit- und Informationsmangel) nicht in der Lage, alle möglichen Handlungsalternativen, alle möglichen Konsequenzen und deren Eintrittswahrscheinlichkeiten abzuschätzen; wir haben oft keine genauen Vorstellungen über den Wert der möglichen Konsequenzen; wir sind oft voreingenommen, so dass wir die Konsequenzen nicht unparteiisch abwägen, sondern bei für uns vorteilhaften Handlungen den Schaden für andere unterschätzen. Außerdem hätten wir, wenn alle nach diesem Entscheidungskriterium handelten, keine Erwartungssicherheit mehr, und die menschliche Kooperation würde zusammenbrechen, da wir uns nicht mehr darauf verlassen könnten, dass Abmachungen, Verträge, Regeln etc. eingehalten werden.

Aufgrund dieser Einwände wird der direkte Handlungsutilitarismus von niemandem ernsthaft vertreten. Stattdessen vertreten Utilitaristen die Auffassung, dass wir nach moralischen Regeln handeln sollen, nämlich idealerweise nach denjenigen moralischen Regeln, deren allgemeine Internalisierung den größten zu erwartenden Gesamtnutzen hat. Da es aber unmöglich ist, herauszufinden, welche Regeln diese Bedingung erfüllen, beschränken sich Utilitaristen auf folgendes Entscheidungskriterium: Handle (von außergewöhnlichen Situa-

11 Vgl. zu den folgenden Einwänden: Dieter Birnbacher, *Analytische Einführung in die Ethik*, 2., durchgesehene und erweiterte Auflage, Berlin 2007, S. 194 ff.

tionen abgesehen) in jeder Situation nach den altbewährten Regeln unserer herkömmlichen Moral.

Einen Handlungsutilitarismus, der dessen Richtigkeitskriterium mit einem Entscheidungskriterium verbindet, das verlangt, nach bestimmten Regeln zu handeln, bezeichnet man als **indirekten Handlungsutilitarismus**: Den besten zu erwartenden Gesamtnutzen erreicht man am ehesten, wenn man ihn nicht direkt intendiert, sondern den indirekten Weg über die Befolgung moralischer Regeln einschlägt: Das Handeln nach Regeln ist das beste Mittel zur Ausführung der Handlung mit dem größten zu erwartenden Gesamtnutzen.

Dieses Mittel erfüllt seinen Zweck aber nur dann, wenn sich die meisten Menschen meistens an die Regeln halten. Deshalb dürfen die Regeln keine bloßen Faustregeln sein, die man als grobe Richtlinien akzeptiert, aber jederzeit ohne schlechtes Gewissen verletzt, falls man glaubt, dadurch einen größeren zu erwartenden Gesamtnutzen zu erzielen. Sie müssen vielmehr internalisierte moralische Regeln sein, die wir nur widerwillig und mit schlechtem Gewissen verletzen.

Kritiker wenden ein, dass es nicht möglich ist, eine solche Haltung gegenüber den Regeln einzunehmen, wenn wir wissen, dass sie nur instrumentellen Wert haben und die moralische Richtigkeit von Handlungen nicht von diesen Regeln (sondern von den Handlungskonsequenzen) abhängt. Auch wird eingewandt, dass sich kein Kriterium dafür angeben lässt, wann man die Regeln verletzen darf bzw. soll.

Trotz dieser Schwierigkeiten vertreten alle Handlungsutilitaristen einen indirekten Handlungsutilitarismus. Eine bekannte Form des indirekten Handlungsutilitarismus ist R. M. Hares Zwei-Ebenen-Theorie des moralischen Denkens.

## 9. Standardevinwände gegen den Utilitarismus

Obwohl die utilitaristischen Ideen, dass es in der Moral nur um das Wohlbefinden geht und wir den Gesamtnutzen maximieren sollen, auf den ersten Blick äußerst einleuchtend sein mögen, gibt es dennoch zahlreiche Gegner des Utilitarismus. Diese wenden ein, dass er praktisch nicht anwendbar sei sowie kontraintuitive Implikationen habe, da er einerseits zu viel erlaube und andererseits zu viel verlange:

I. Der Utilitarismus ist praktisch nicht anwendbar: Wir können weder wissen, welche Handlung die besten zu erwartenden Konsequenzen hat, da wir unmöglich alle Handlungsalternativen mit allen ihren möglichen Konsequenzen und Eintrittswahrscheinlichkeiten bestimmen können, noch können wir nach der Ausführung einer Handlung wissen, welche tatsächlichen Konsequenzen sie hat, da wir ihre in ferner Zukunft liegenden Konsequenzen nicht kennen können. Folglich kann man nie wissen, welche Handlung moralisch richtig ist. Ein weiterer Kritikpunkt lautet, dass der zur Berechnung des Gesamtnutzens notwendige interpersonelle Nutzenvergleich, also ein Vergleich des Nutzens, den verschiedene Personen aus einer Handlung ziehen, nicht möglich ist.

II. Der Utilitarismus ist, so die Kritiker, mit den wohlüberlegten Moralurteilen unserer herkömmlichen Moral unvereinbar, da er in zweierlei Hinsicht kontraintuitiv ist:

1. Aufgrund des Maximierungsgebots fordert der Utilitarismus zu viel und überfordert die Menschen: Er unterscheidet nicht zwischen moralisch gebotenen und supererogatorischen Handlungen (also lobenswerten Handlungen, die nicht geboten sind) und gebietet Handlungen, die von niemandem moralisch gefordert werden können, weil sie zu große Opfer verlan-

gen. Er schränkt unsere Handlungsfreiheit auf unzumutbare Weise ein und macht es unmöglich, ein gutes Leben zu führen, da wir fast alle persönlichen Ziele, Projekte und Vergnügungen aufgeben müssen, um stattdessen unser Geld, unsere Zeit und Arbeitskraft der Linderung der Übel dieser Welt zu widmen. Eine Reaktion auf diesen Überforderungseinwand ist der (zuerst von Michael Slote vertretene) **satisfizierende Utilitarismus** (*satisficing utilitarianism*), in dem das Maximierungsgebot aufgegeben wird: Die Handlung mit dem größten Gesamtnutzen ist immer erlaubt, geboten ist jedoch nur die Handlung, deren Gesamtnutzen groß genug ist. Neben der offensichtlichen Schwierigkeit, zu bestimmen, wann der Gesamtnutzen groß genug ist, hat der satisfizierende Utilitarismus zahlreiche kontraintuitive Implikationen und wird inzwischen kaum noch vertreten.<sup>12</sup>

2. Da der Utilitarismus nur die Konsequenzen von Handlungen als moralisch relevant anerkennt, berücksichtigt er, so die Kritiker weiter, viele Faktoren gar nicht, die wir in unserer herkömmlichen Moral für moralisch relevant halten. Dies führt dazu, dass der Utilitarismus zu viel erlaubt, indem er um des größten Gesamtnutzens willen Handlungen erlaubt (und sogar gebietet), die gemäß der herkömmlichen Moral offensichtlich verboten sind:

- (a) Er ignoriert die intrinsische Falschheit bzw. Schlechtheit bestimmter Handlungsweisen (wie das absichtliche Töten Unschuldiger) und erlaubt solche Handlungsweisen schon zur Erzielung geringfügig besserer Konsequenzen.
- (b) Er lässt alle Freuden und Präferenzen bei der Berechnung

12 Vgl. zum satisfizierenden Utilitarismus Jason Rogers, In Defense of a Version of Satisficing Consequentialism, *Utilitas* 22 (2010), S. 199–221; Ben Bradley, Against Satisficing Consequentialism, *Utilitas* 18 (2006), S. 97–108.

der besten Konsequenzen zu und ignoriert, dass bestimmte Freuden und Präferenzen (z. B. die sadistische Freude am Leid anderer oder die Freude bei einer Vergewaltigung) unmoralisch sind und keinerlei Anspruch auf Berücksichtigung haben.

- (c) Er ignoriert aufgrund der geforderten Unparteilichkeit besondere Beziehungen zwischen Personen und daraus resultierende spezielle Rechte und Pflichten (zwischen Freunden, Eltern und Kindern etc.): Gesetzt den Fall, dass wir nur eines von zwei Kindern retten könnten, gebietet uns die herkömmliche Moral, unser eigenes Kind einem fremden Kind vorzuziehen, während im Utilitarismus eine solche Parteilichkeit verboten wäre.
- (d) Er respektiert keine moralischen Rechte von Personen.
- (e) Er ignoriert die Würde des Menschen und erlaubt, Menschen als bloße Mittel zum Zweck zu instrumentalisieren, indem er die Bestrafung Unschuldiger sowie das Opfern von Menschen zur Rettung einer größeren Anzahl anderer Menschen erlaubt.
- (f) Er ignoriert die Bedeutung vergangener Handlungen und daraus resultierende spezielle Rechte und Pflichten (wie z. B. Dankbarkeit und das Einhalten von Versprechen).
- (g) Er ignoriert Verteilungsgerechtigkeit und beurteilt Zustände nur nach der Höhe des Gesamtnutzens, also unabhängig davon, wie dieser Nutzen auf einzelne Individuen verteilt ist. Verteilungsgerechtigkeit wird nur berücksichtigt, insofern sie aufgrund des abnehmenden Grenznutzens bestimmter Güter zur Maximierung des Gesamtnutzens beiträgt: Besitzt man einen Wintermantel, verringert sich der Nutzen jedes weiteren Wintermantels, so dass der Gesamtnutzen größer ist, wenn vier Personen je einen Wintermantel besitzen, als wenn eine Person vier Wintermäntel besitzt.

Die Überforderungs- und Kontraintuitivitätseinwände können in dem (in diesem Band abgedruckten, vgl. S. 244) Vorwurf von John Rawls zusammengefasst werden, dass der Utilitarismus die ›Getrenntheit der Personen‹ nicht beachtet: Es ist zwar klug, wenn eine Person Lasten auf sich nimmt, um später davon zu profitieren, aber es ist unmoralisch, dieses für eine Person kluge Verhalten auf mehrere Personen zu übertragen und einigen Personen Lasten aufzubürden, von denen andere Personen profitieren.

Weiterhin bestreitet der Utilitarismus, so die Kritiker, die moralische Relevanz einiger Unterscheidungen unserer herkömmlichen Moral: Er bestreitet die moralische Relevanz der Unterscheidung zwischen Tun und Unterlassen bzw. Zulassen. Dies hat zur Folge, dass wir für Unterlassungen im selben Maße moralisch verantwortlich sind wie für unsere Handlungen und dass wir für jedes Übel, das wir nicht lindern oder verhindern, obwohl wir es könnten, moralisch verantwortlich sind. Diese überbordende »negative Verantwortung« wird in diesem Band von Bernard Williams kritisiert (vgl. S. 271 ff.).

Der Utilitarismus bestreitet auch die moralische Relevanz der Unterscheidung zwischen beabsichtigten Folgen und vorausgesehenen, aber unbeabsichtigten Nebenfolgen. Wir sind daher einerseits für die unbeabsichtigten Nebenfolgen unserer Handlungen ebenso moralisch verantwortlich wie für die beabsichtigten Folgen und dürfen andererseits (um des größeren Gesamtnutzens willen) bestimmte Folgen (wie z. B. den Tod unschuldiger Zivilisten im Krieg) absichtlich herbeiführen, die man gemäß weitverbreiteten moralischen Überzeugungen nur als unbeabsichtigte Nebenfolgen in Kauf nehmen, aber nicht direkt beabsichtigen darf.

Der Einwand der Kontraintuitivität des Utilitarismus wird meistens begründet, indem man fiktive Beispielfälle konstruiert und behauptet, Utilitaristen müssten in diesen Fällen of-

fensichtlich falsche Handlungen für moralisch richtig erklären. Ein bekanntes Beispiel dieser Art lautet:<sup>13</sup> In einem Krankenhaus liegen fünf Patienten, deren Leben nur dann gerettet werden kann, wenn ihnen unterschiedliche Organe transplantiert werden. Zufällig ist gerade ein alleinstehender Mann (der keine Freunde hat) zu einer Routineuntersuchung im Krankenhaus, der als Organspender für alle fünf Patienten geeignet ist. Utilitaristen müssten in diesem Fall behaupten, die Ärzte sollen unauffällig den Tod des Mannes arrangieren, um mit dessen Organen die anderen fünf Patienten zu retten, da ein Toter besser ist als fünf Tote. Utilitaristen gebieten in diesem Fall also eine Handlung, die klarerweise moralisch falsch ist. (Dieses Beispiel illustriert gleichzeitig den Überforderungseinwand, da der alleinstehende Mann sogar die Pflicht hätte, sich für die anderen fünf Patienten zu opfern.)

Beispiele dieser Art weisen Utilitaristen mit folgenden Argumenten zurück:

1. Die Beispiele sind zu einfach und unrealistisch und ignorieren viele Unwägbarkeiten und langfristige Konsequenzen. In der Realität werden Situationen, in denen die intuitiv falsche Handlung den größten Gesamtnutzen hat, niemals auftreten.
2. In den Beispielen wird den Akteuren ein Wissen um die Umstände und Konsequenzen zugeschrieben, das diese nicht haben können. Akteure sollten daher und aufgrund der oben genannten Einwände gegen das Entscheidungskriterium des direkten Handlungsutilitarismus nicht versuchen, den größten Gesamtnutzen zu ermitteln, sondern sich (gemäß dem indirekten Handlungsutilitarismus) an mora-

13 Weitere instruktive Beispiele finden sich in Tim Mulgan, *Understanding Utilitarianism*, Stocksfield 2007, S. 93 ff.