KRITERION, Nr.15 (2002), pp.13–27

Fabian Fricke

VERSCHIEDENE VERSIONEN DES NEGATIVEN UTILITARISMUS

Abstract: Karl Popper suggested that the utilitarian formula "Maximise happiness" should be replaced by the formula "Minimize suffering" – a view which has been called "negative utilitarianism". Can such a view be spelled out in a plausible way? Examining several ways of understanding the alleged ethical priority of pain over pleasure, I come to the conclusion that none of them is satisfactory as a basis for a theory of benevolence. However, once we move from hedonistic to preference utilitarianism two possible views emerge which, although not fundamentally different from classical utilitarianism, diverge from it in interesting ways and seem to warrant further study.

Karl Popper hat mit einigen Bemerkungen in den Fußnoten seines Buches *The Open Society and Its Enemies* eine ethische Theorie ins Leben gerufen, der ein nur mäßig glorreiches Schicksal beschert war: den negativen Utilitarismus. Obwohl die Überlegungen, die ihn zu der von ihm vorgeschlagenen Modifikation des utilitaristischen Prinzips führten, durchaus auch auf andere Denker einen großen Reiz ausgeübt haben, scheinen einige schwerwiegende Kritikpunkte seinen Vorschlag als letztlich unbrauchbar erwiesen zu haben, und in heutigen Darstellungen des Utilitarismus sieht sich der negative Utilitarismus häufig erneut in die Fußnoten verbannt (z. B. Shaw, 1999, S. 233n.). Dass dies nicht völlig zu Recht geschieht und dass Poppers Intuitionen durchaus noch wertvoll für uns sein können, soll in diesem Aufsatz gezeigt werden.

Worum geht es? Popper behauptet, dass moralisch gesehen keine Symmetrie zwischen Glück und Leid besteht: "We should realize that from the moral point of view suffering and happiness must not be treated as symmetrical; that is to say, the promotion of happiness is in any case much less urgent than the rendering of help to those who suffer, and the attempt to prevent suffering." (Popper, 1945, I 5 n. 6) Er rät daher, die utilitaristische Formel "Maximiere Glück!" durch die Formel "Minimiere Leid!" zu ersetzen. Popper selbst war kein Utilitarist und hat in dieser Formel nur ein plausibles Prinzip gesehen, das neben anderen zu einer Grundlage öffentlicher Politik werden könne, doch liegt natürlich die Frage nahe (und ist gestellt worden), ob auch reine Utilitaristen ihre Theorie in der vorgeschlagenen Weise modifizieren sollten. Allgemeiner gesprochen könnte man sich fragen, ob das poppersche Prinzip eine adäquate Grundlage für denjenigen Bereich der Moral darstellen würde, den man als Theorie des unparteiischen Wohlwollens bezeichnen könnte: Selbst wenn man nicht wie die Utilitaristen glaubt, dass diese die gesamte Moral ausmacht, enthalten doch viele pluralistische Ethiken so etwas wie ein Prinzip des Wohlwollens, dessen Ausformulierung in einer gewissen Unabhängigkeit zum Rest des Systems geschehen kann.

Die Frage, inwieweit der negative Utilitarismus eine plausible Theorie des Wohlwollens darstellt, dürfte also von relativ breitem Interesse sein. Dazu wäre aber noch näher zu klären, wie genau der negative Utilitarismus auszusehen hat; tatsächlich scheint der poppersche Vorschlag eher eine Familie von verwandten Theorien anzuregen als selbst schon eine komplette Theorie zu liefern. Dieser Aufsatz soll eine Übersicht und Beurteilung der verschiedenen Gestalten liefern, die der negative Utilitarismus annehmen kann.

I

Der negative Utilitarismus, so wie er gewöhnlich verstanden wird, behauptet also die Existenz einer Asymmetrie zwischen Glück und Leid in moralischer Hinsicht; die Förderung von Glück sei "weniger dringend" als die Linderung und Verhinderung von Leid. Es gibt aber natürlich verschiedene Möglichkeiten, wie die moralische Priorität des Leides aufgefasst werden kann, und entsprechend ergeben sich verschiedene Versionen eines negativen Utilitarismus. Schauen wir uns einige Möglichkeiten an:

1) praktischer Vorrang der Leidminderung
2) relativer Vorrang der Leidminderung
3) absoluter Vorrang der Leidminderung
   a) Glücksförderung obligatorisch
   b) Glücksförderung supererogatorisch
   c) Glücksförderung moralisch neutral

Nach der ersten Version besteht insofern eine Asymmetrie zwischen Glück und Leid, als es in der Praxis normalerweise viel effizienter ist, Quellen des Leides zu eliminieren, als sich um die Erzeugung von Glück zu kümmern. Es ist nicht wenig wahrscheinlich, dass wir mehr für andere Wesen tun können, indem wir sie von ihrem Leid befreien, als indem wir nach Wegen suchen, ihr Glück zu vergrößern, und ein Utilitarist wäre wahrscheinlich gut beraten, "Minimiere Leid!" zu seiner ersten praktischen Maxime zu machen.

Poppers Behauptung, die Förderung von Glück sei "in any case much less urgent" als die Verminderung von Leid, mag darauf hindeuten, dass derartige Überlegungen auch bei ihm am Anfang standen; sie reichen aber nicht aus, um den negativen Utilitarismus zu einer vom klassischen Utilitarismus verschiedenen Theorie zu machen. Sie haben ihren Platz vielmehr innerhalb des klassischen Utilitarismus: Der Utilitarist wird zugestehen, dass es normalerweise besser ist, der Verminderung von Leid den Vorrang zu geben, wird aber darauf beharren, dass es Ausnahmefälle geben wird, in denen es moralisch richtig ist, das Glück eines Wesens zu vermehren, auch wenn ein anderes dadurch ein vergleichsweise geringes Leid erfahren wird. Wenn es dem negativen Utilitarismus nur um die praktische Effizienz ginge, dann handelte es sich bei ihm um nicht mehr als um eine Mahnung an die Utilitaristen, dem Leid die ihm empirisch zukommende Beachtung zu schenken. Dies ist in der Tat der Beitrag zur Theorie des Utilitarismus, der dem negativen Utilitarismus oft zuerkannt wurde (z. B. J. J. C. Smart, 1973, S. 29f.; Scarre, 1996, S. 17f.).

So verstanden wäre die Asymmetrie zwischen Glück und Leid also allein auf der praktischen, nicht aber auf der theoretischen Ebene angesiedelt. Für viele wird die Anziehungskraft des negativen Utilitarismus aber eben gerade darin bestehen, dass dem Leid schon auf theoretischer Ebene eine höhere Bedeutung zugemessen wird als dem Glück: dass unser moralisches Verhältnis zu beiden verschieden ist. Denkbar wäre nun, dass der Vorrang des Leides ein bloß relativer wäre, oder er könnte sogar ein absoluter sein.

Die Idee eines relativen Vorranges des Leides mag attraktiv erscheinen, da sie gemäßigter auftritt; sie leidet jedoch unter theoretischen Schwächen. Sie würde besagen, dass jedes Leid nur durch eine deutlich größere Menge Glück aufgewogen würde, indem wir etwa im hedonistischen Kalkül alle Leidwerte mit einem bestimmten Faktor multiplizierten. Dies scheint jedoch vorauszusetzen, dass wir eine sinnvolle Skala aufstellen können, anhand derer wir von einander entsprechenden Mengen an Glück und Leid sprechen können. Eine solche Skala ließe sich aber wohl nur aufstellen, indem wir von vergleichenden Präferenzurteilen über bestimmte glück- und leidvolle Erfahrungen ausgehen – es scheint kaum möglich zu sein, Glück und Leid unabhängig von solchen Urteilen auf einer gemeinsamen Skala zu

messen (cf. Griffin, 1986, S. 84). Wir müssten also bewusst diese Präferenzskala um einen bestimmten Faktor verzerren, und es ist schwer einzusehen, woher wir die Rechtfertigung dafür nehmen sollten. Wenn jemand tatsächlich der Ansicht ist, eine bestimmte Menge Glück würde eine bestimmte Menge Leid aufwiegen, und lieber beides erfahren würde als keines von beidem – auf welcher Basis sollten wir geltend machen, in Wahrheit würde erst die x-fache Menge an Glück dieses Leid aufwiegen? Wenn hingegen nie jemand eine noch so große Menge Glück für ausreichend erachten würde, um ein noch so kleines Leid zu kompensieren, dann wäre es unnötig, dem Leid zusätzliches Gewicht zu verleihen – eine angemessene Werteskala würde dies bereits reflektieren.

Es mag der Eindruck entstehen, dass diese Überlegungen auch der Idee eines absoluten Vorranges des Leides und damit wohl dem negativen Utilitarismus überhaupt den Todesstoß versetzen dürften: Dieser scheint vorauszusetzen, dass unser Leid nie durch noch so große Mengen an Glück aufgewogen werden kann. Wenn dies (was wahrscheinlich erscheint) unseren tatsächlichen Präferenzen widerspricht, kann dann ein absoluter Vorrang der Leidminderung überhaupt noch begründet werden?

Theoretisch schon, denn es ist zwar plausibel, aber keineswegs zwingend, dass unsere tatsächlichen Präferenzen den wirklichen Wert der verschiedenen Optionen bestimmen. Der Gedanke, wir könnten uns in unseren gewöhnlichen Präferenzen irren, und in Wahrheit könne Leid nie durch Glück aufgewogen werden, mag unseren Intuitionen widersprechen, ist aber verständlich und näherer Betrachtung wert. Hingegen ist die für die Begründung eines relativen Vorranges notwendige Idee, Leid könne jeweils nur durch die x-fache Menge Glück aufgewogen werden, als dies unsere Präferenzen nahe legen, von vornherein von geringer Überzeugungskraft – dies wäre eine höchst merkwürdige Art von Irrtum.

Bevor wir die Plausibilität eines absoluten Vorranges näher prüfen, soll noch kurz auf die drei Varianten eingegangen werden, die ich oben aufgeführt habe. Sie stellen drei mögliche Versionen eines negativen Utilitarismus dar und betreffen weniger das Verhältnis von Glück und Leid als vielmehr weitere Vorstellungen, wie die Theorie des Wohlwollens auszusehen hat. Die erste Variante geht von einer lexikalischen Ordnung zwischen Glück und Leid aus: Die erste Pflicht jedes Handelnden bestünde darin, das Leid zu minimieren; ist dies erreicht, und sind mehrere Alternativen hinsichtlich des Leides gleichwertig, gilt es das Glück zu maximieren.

Schon Popper hatte in der "Bescheidenheit" der Forderung "Minimiere Leid!" einen Vorteil des negativen Utilitarismus gesehen, und die Tatsache, dass dieser offenbar geringere Forderungen an den Handelnden stellt als der oft der Überforderung des Handelnden beschuldigte klassische Utilitarismus, stellte auch für spätere Philosophen einen Anziehungspunkt dieser Theorie dar (z. B. Walker, 1974, S. 425). Dies gibt dieser Ansatz auf. Umgekehrt empfiehlt er sich aber für Konsequentialisten, die in der Forderung, das Glück anderer Wesen zu maximieren, nichts Anstößiges sehen, am klassischen Utilitarismus jedoch kritisieren, dass er es gebietet, einem Wesen Leid zuzufügen, um das Glück eines anderen zu vergrößern. Diese Konsequentialisten könnten an der ersten Variante des negativen Utilitarismus Interesse zeigen.

Die zweite Variante erklärt dagegen die Förderung des Glückes für supererogatorisch: Wenn wir nach unseren Kräften das Leid der Welt minimiert haben, haben wir unsere Pflicht erfüllt, und alle darüber hinausgehenden Dienste an anderen Wesen sind lobenswert, aber nicht obligatorisch.

Diese Variante bietet sich offensichtlich für all jene an, welche die Überforderung des Handelnden im klassischen Utilitarismus für ein Problem halten und eine Theorie des Wohlwollens vorziehen, die sich mehr in Richtung Commonsense bewegt.

Nach der dritten Variante schließlich wäre die Förderung des Glückes als moralisch neutral zu betrachten. Dies mag intuitiv nicht sehr plausibel erscheinen, hat aber den Vorteil theoretischer Strenge: Diese Version des negativen Utilitarismus kommt mit einem einzigen, negativen Wert aus, den es zu minimieren gilt, und könnte so ihren Reiz haben für minimalistisch eingestellte Konsequentialisten, welche die Überforderung des Handelnden scheuen, aber die Idee der Supererogation für theoretisch suspekt halten.

Wir brauchen hier die relativen Verdienste der drei Varianten nicht zu diskutieren; sie sollen nur zeigen, dass der negative Utilitarismus für verschiedene Gruppen attraktiv sein könnte, wenn denn die Grundidee eines absoluten Vorranges des Leides über das Glück plausibel gemacht werden könnte, und dieser Frage werden wir uns jetzt zuwenden. Dazu gilt es zuerst zu definieren, was überhaupt unter "Glück" und "Leid" zu verstehen sei. Diese Unterscheidung gehört in erster Linie der hedonistischen Spielart des Utilitarismus an, und entsprechend sollten wir die Begriffe am besten über die Qualität mentaler Zustände definieren: Als Leid hat jeder mentale Zustand zu gelten, dem gegenüber wir ablehnend eingestellt sind, bei dem wir es also vorziehen würden, ihn nicht erleben zu müssen. Umgekehrt soll unter "Glück" jeder mentale Zustand verstanden werden, dem gegenüber wir mindestens neutral eingestellt sind, gegen dessen Erleben wir uns normalerweise nicht wehren würden.

Die These des negativen Utilitarismus kann nun darin gesehen werden, dass kein Zustand des Leides je durch einen Zustand des Glückes kompensiert werden kann, und so finden wir sie auch bei Popper ausgedrückt: "But, from the moral point of view, pain cannot be outweighed by pleasure, and especially not one man's pain by another man's pleasure." (Popper, 1945, I 9 n. 2)

Popper betont, dass dies vom "moral point of view" aus gelte; wie sollen wir das verstehen? Offenbar doch so, dass es nie möglich ist, ein Leid, das man einem anderen verursacht, durch eine noch so große Menge Glück entweder für dasselbe Wesen oder für andere Wesen zu rechtfertigen. Lassen wir den zweiten Fall erst einmal beiseite und konzentrieren wir uns auf den ersten: Kann kein Leid, das mir zugefügt wird, je durch eine noch so große Menge Glück kompensiert werden?

Dies scheint sehr fraglich, zumindest wenn es um von uns selbst gewähltes Leid geht: Die meisten Menschen scheinen durchaus bereit, freiwillig für eine kurze Zeit einen mentalen Zustand auszuhalten, den sie lieber nicht erleben würden, um dadurch ihre Glücksmenge zu mehren (nicht nur, um späteres Leid zu verhindern): Schönheitsoperationen, Tätowierungen, manche Formen sportlicher Betätigung, Mehrarbeit, um sich Luxusgüter (statt nur dem Lebensnotwendigen) leisten zu können – dies sind offensichtliche Beispiele dafür.

Die Dinge werden noch deutlicher, wenn wir uns den Bereichen zuwenden, in denen der negative Utilitarismus traditionell (und ziemlich vernichtend) kritisiert wurde: Zeugung und Tod (cf. R. N. Smart, 1958, S. 542f.). Wenn wir davon ausgehen, dass ein Wesen in seinem Leben noch mindestens einen Augenblick erleben wird, den es lieber nicht durchleben würde, dann wäre es vom Standpunkt der reinen Leidminimierung aus am besten, das Wesen schmerzlos zu töten. Notfalls wäre auch ein mehr oder weniger schmerzvoller Tod zu rechtfertigen – wenn wir etwa annehmen müssten, jemand würde in der Zukunft einige Tage an Zahnschmerzen leiden müssen, wäre es sicher besser für ihn,

ihn rasch zu erstechen. Offensichtlich dürften solche Bedingungen auf so ziemlich alle Menschen zutreffen, und so wäre es zu ihrem eigenen Wohl geboten, sie so schnell und schmerzlos wie möglich zu töten. Ebenso klar wäre, dass es ein Verbrechen gegen ein ungezeugtes Kind wäre, es in die Welt zu setzen, da sein Leben mit Sicherheit Augenblicke enthalten wird, die es lieber nicht erleben würde.

Zwar haben Pessimisten aller Zeiten ähnliche Gedanken geäußert, doch scheint eine solche Einstellung dem Leben gegenüber wenig mehrheitsfähig. Die Idee, wir müssten unserem Mörder danken, unseren Lebensretter aber verfluchen, mag nicht völlig absurd sein, sie ist aber für eine allgemein annehmbare Theorie des Wohlwollens kaum zu gebrauchen. Wenn dies der Kern des negativen Utilitarismus ist, dann stellt er in der Tat keine ernsthaft erwägenswerte Alternative zu anderen Theorien dar.

So schnell sollten wir uns jedoch nicht von ihm verabschieden, denn vielleicht ließen sich Modifikationen denken, welche die Grundidee des negativen Utilitarismus bewahren können, aber diese extremen Konsequenzen vermeiden. Insbesondere die folgende springt ins Auge: Nach dem obigen Popper-Zitat kann Leid nie durch Glück kompensiert werden, weder durch eigenes noch durch das anderer Wesen. Ersteres scheint mit guten Gründen bezweifelbar zu sein, aber wie sieht es mit Letzterem aus? Können wir nicht sagen, dass eigenes Leid zwar durch eigenes Glück kompensiert werden kann, nicht aber durch fremdes Glück? Dies ist ja ein Standardeinwand gegen den Utilitarismus; können wir den negativen Utilitarismus als einen Versuch verstehen, diesem Einwand Rechnung zu tragen? Versuchen wir, eine entsprechende Theorie zu skizzieren.

Damit die obigen, unangenehmen Konsequenzen vermieden werden können, darf nicht mehr die Minimierung des Leides überhaupt, sondern nur diejenige des unkompensierten Leides gefordert sein. Es stellen sich zwei Fragen: Unter welchen Umständen gilt Leid als kompensiert? Gibt es auch Einschränkungen, die das Verursachen oder Zulassen von kompensiertem Leid betreffen?

Es sind zwei Arten von Kompensation denkbar, die man als synchrone und diachrone Kompensation bezeichnen könnte. Synchrone Kompensation findet statt, wenn ein als unangenehm empfundener mentaler Zustand durch einen gleichzeitigen angenehmen mentalen Zustand mehr als ausgeglichen wird: Während ich ein gutes Buch lese, kann ich einen leichten Kopfschmerz verspüren; verspürte ich nur diesen, würde ich die fragliche Zeit lieber nicht bewusst erleben, aber die Kombination mit der Lektüre des Buches lässt das Erleben meines gesamten mentalen Zustandes als lohnend erscheinen. Bei der diachronen Kompensation dagegen folgt das angenehme Erlebnis dem Leid oder geht ihm voraus: So renne ich etwa zehn Minuten durch Kälte und Regen, um mir einen Film im Kino anzuschauen, oder ich betrinke mich, obwohl ich nachher unter einem Kater leiden werde.

Die letzten Beispiele zeigen bereits, dass wir gewöhnlich auch die diachrone Kompensation als echte Kompensation anerkennen. In den meisten Fällen liegen dabei Glück und Leid zeitlich eng beieinander und stehen in einer engen kausalen Verbindung, doch scheint dies nicht notwendig zu sein: Wenn uns das Verhältnis von Glück und Leid angemessen erscheint, dann würden wir wohl auch jetzt ein geringes Leid auf uns nehmen, wenn wir mit ausreichender Sicherheit in ferner Zukunft eine entsprechend größere Menge Glück erleben werden (oder umgekehrt) – und wir dürften auch nicht allzu sehr zögern, das Gleiche einem anderen Wesen aufzuerlegen, wenn wir von seiner mutmaßlichen Zustimmung ausgehen können.

Dies hieße aber, dass wir es nur noch dann mit unkompensiertem Leid zu tun haben, wenn der gesamte Rest des Lebens des betroffenen Wesens von ihm selbst als nicht mehr erlebenswert angesehen würde. Das hätte die erwünschte Konsequenz, dass wir schlimmstenfalls noch die Existenz solcher bedauernswerter Wesen zu beenden hätten (und dabei geht es um Fälle, wo auch der Commonsense eine Euthanasie in Erwägung ziehen würde), nicht aber die all jener Wesen, die ihrem Weiterleben noch einen positiven Wert zuschreiben würden. Das größte Problem des negativen Utilitarismus wäre also vermieden.

Andererseits würde ein solches Gebot der Minimierung unkompensierten Leides allein uns eine zu freie Hand bei der Behandlung derjenigen Wesen lassen, in deren Leben das Glück das Leid überwiegt: Wird unsere Theorie des Wohlwollens nicht ergänzt, dürften wir ihnen beliebiges Leid zufügen oder geschehen lassen, solange die Bilanz für jedes einzelne Wesen immer noch knapp positiv wäre. Haben wir die Menge des unkompensierten Leides nach Kräften minimiert, stehen uns wie oben drei Möglichkeiten offen, wie wir das mit dem restlichen Leid verrechnete Glück betrachten sollen: als gebotenermaßen zu maximieren, als Gut, dessen Förderung zwar lobenswert, aber nicht geboten ist, oder als moralisch neutral. Hier scheint nun (sofern wir nicht irgendwelche deontologischen Verbote in unsere Theorie aufnehmen wollen) einzig die erste Variante einen unattraktiven moralischen Minimalismus zu vermeiden: Dass es moralisch neutral oder zumindest erlaubt sein soll, anderen Wesen ein beliebiges Maß an Leid zuzufügen, solange dies von ihrem Glück über ihr ganzes Leben hinweg immer noch kompensiert wird, erscheint kaum annehmbar.

Was sollen wir nun von unserer derartig modifizierten Theorie – dem negativen Utilitarismus mit Kompensation – halten? Betrachten wir einen Fall, in dem wir es sowohl mit kompensiertem als auch mit unkompensiertem Leid zu tun haben:

	w_1	w_2	w_3
x_1	2/-5	4/-6	2/-6
x_2	3/-1	3/-1	6/-1

x_1 und x_2 sind zwei Individuen, w_1 bis w_3 drei mögliche Welten. Die Zahlen vor und hinter den Schrägstrichen stehen für die Mengen an Glück (positiv) und Leid (negativ), die x_1 und x_2 in den möglichen Welten jeweils erleben würden.

Für den klassischen Utilitarismus wäre w_3 die beste Welt: Wir addieren einfach die Werte für das Glück, ziehen diejenigen für das Leid ab und erhalten so folgende Gesamtbewertung:

	w_1	w_2	w_3
x_1	-3	-2	-4
x_2	2	2	5
x_1+x_2	-1	0	1

Zwar ist x_1 in w_3 am schlechtesten dran, doch wird dies durch die höhere Glücksmenge von x_2 kompensiert. Dies ist natürlich genau die Art von Ergebnis, gegen das sich der negative Utilitarismus wendet. Dieser würde nur die negativen Werte betrachten und käme so zu folgender Gesamtbewertung:

	w_1	w_2	w_3
x_1	-5	-6	-6
x_2	-1	-1	-1
x_1+x_2	-6	-7	-7

Für den negativen Utilitarismus würde also die vorher als schlechteste eingeschätzte Welt w_1 zur besten. Aber wie wir gesehen haben, kann man auch daran Zweifel hegen: Zwar hat x_1 in w_2 mehr zu leiden als in w_1, doch wird dies durch eine größere Glücksmenge für x_1 selbst mehr als kompensiert (x_2 verhält sich neutral

zwischen w_1 und w_2). Hier kommt nun der negative Utilitarismus mit Kompensation zum Tragen:

	w_1	w_2	w_3
x_1	-3	-2	-4
x_2	0	0	0
x_1+x_2	-3	-2	-4

Indem wir für jedes Individuum Glück und Leid einzeln miteinander verrechnen und dann nur die Fälle betrachten, in denen das Ergebnis negativ ist, wird w_2 zur besten Welt. Und dies ist intuitiv durchaus nachvollziehbar: Für x_1 ist w_2 gesamthaft gesehen das beste Ergebnis; für x_2 gäbe es mit w_3 ein besseres, doch da dieses mit unkompensiertem Leid für x_1 erkauft würde und es x_2 in w_2 immer noch gut geht, stellt dies keinen intuitiv zwingenden Grund dar, w_3 vorzuziehen.

Eingefleischte Utilitaristen werden zwar auch in unserem Beispiel dem Verdikt der klassischen Theorie den Vorzug geben, doch scheint der negative Utilitarismus mit Kompensation hier eine durchaus attraktive Alternative zu bieten. Es muss jedoch daran erinnert werden, dass dies nur in all jenen Fällen gilt, in denen wir es tatsächlich mit unkompensiertem Leid zu tun haben, also mit Wesen, deren zukünftiges Leben von ihnen als nicht erlebenswert beurteilt würde. Derartige Situationen dürften nicht allzu häufig sein, und in anderen Fällen treffen sich die Verdikte unserer Theorie mit denjenigen des klassischen Utilitarismus. Dennoch hat sie immerhin ein ganz brauchbar erscheinendes Bollwerk gegen die schlimmsten der angenommenen Exzesse des Utilitarismus aufgebaut: Wir können ein Wesen zwingen, zugunsten des größeren Glückes anderer Lasten zu tragen – solange dies sein Leben nicht vollständig ruinieren würde. Dies könnte man als einen akzeptablen Kompromiss zwischen Vertretern des Utilitarismus und seinen Kritikern ansehen.

Es bleibt jedoch ein Problem: Im Gegensatz zum gewöhnlichen negativen Utilitarismus schätzt unsere Theorie ein normales, erfülltes Leben, das seine Höhen und Tiefen aufweist, zwar als ein Gut ein – aber als ein Gut von relativ geringer Bedeutung, und daher kommen wir bei der Frage des Tötens erneut zu Urteilen, die von unserer alltäglichen Beurteilung stark abweichen. Es ist zwar nicht mehr moralisch geboten, das Leben aller normalen Menschen so schmerzlos wie möglich zu beenden, und es ist unter normalen Umständen kein Verbrechen mehr, ein Kind in die Welt zu setzen. Wir wären aber zur Tötung von glücklichen Wesen aufgefordert, wenn dies der einzige Weg wäre, unkompensiertes Leid zu verhindern. Um einem Wesen, dessen Leben im Wesentlichen eine Qual ist, sein Los auch nur ein kleines bisschen zu erleichtern, muss ich notfalls das Leben einer beliebigen Zahl von Wesen opfern, die sonst lange glücklich leben könnten.

Der negative Utilitarismus mit Kompensation fordert uns auf, die Absenz von Glück als unvergleichlich weniger schlimm einzuschätzen als die Existenz von unkompensiertem Leid. Da der Tod nur die Absenz von Glück darstellt, ist er nur ein geringes Übel. Ich würde dies nicht für eine völlig absurde Position halten, und sie hat eine lange philosophische Tradition hinter sich. Diese konzentrierte sich freilich meist auf die Einstellung, die wir zu unserem *eigenen* Tod haben sollten – die Übertragung dieser Einstellung auf den Tod anderer Wesen ist dann zwar eigentlich nur konsequent, aber eher ungewöhnlich und wenig im Einklang mit unseren gewöhnlichen Empfindungen. Unsere modifizierte Version des negativen Utilitarismus mag für einige Leute attraktiv sein, und ich würde sie nicht von vornherein verwerfen – wenn der negative Utilitarist aber (wie Popper) nach einer konsensfähigeren und intuitiv akzeptableren Alternative zum

klassischen Utilitarismus sucht, dann ist er fehlgegangen.

II

Soweit wir bis jetzt gesehen haben, scheint der negative Utilitarismus daran zu scheitern, dass er den Tod nicht als ein angemessen großes Übel einschätzt. Diese Einschätzung muss sich ändern, wenn wir eine brauchbare Theorie des Wohlwollens erhalten wollen – diese darf das Töten oder Sterbenlassen eines glücklichen Wesens nicht als trivial ansehen. Was aber kann an einem schmerzlosen Tod so schlimm sein? Wenn wir von einer möglichen unerfreulichen Existenz nach dem Tode absehen, doch hauptsächlich die Tatsache, dass einem durch den Tod Erlebnisse entgehen, die man noch zu durchleben wünscht. Es ist die Absenz eines Gutes, die wir als ein Übel ansehen – und zwar als ein Übel, das wir offenbar durchaus sogar gegen unkompensiertes Leid abzuwägen für richtig halten. Wir scheinen doch der Ansicht zu sein, dass wir geringes, auch unkompensiertes Leid eines Einzelnen gegen das große Gut der Vielen ausspielen dürfen.

Die Idee des absoluten Vorranges der Leidminderung vor der Glücksförderung (oder Glücksbewahrung) scheint sich also nicht halten zu können. Ist damit der negative Utilitarismus, dessen Kerngedanke die moralische Asymmetrie zwischen Glück und Leid war, endgültig erledigt? Nicht ganz. Der negative Utilitarismus enthält nämlich noch einen zweiten, fundamentaleren Kerngedanken, dem wir uns jetzt zuwenden müssen. Popper drückt ihn in folgendem Satz aus: "It adds to clarity in the field of ethics if we formulate our demands negatively, i.e. if we demand the elimination of suffering rather than the promotion of happiness." (Popper, 1945, I 9 n. 2)

Der negative Utilitarismus legt uns nahe, moralische Forderungen *negativ* zu formulieren: In der Ethik geht es um das Vermeiden von Übeln. Er steht so in einem Kontrast zu einem positiven Utilitarismus, nach welchem es ein moralisches Gut gibt, das es zu fördern gilt. Es ist jedoch nicht gesagt, dass wir nun hedonistisch dieses Gut mit dem Glück und das Übel mit dem Leid identifizieren müssen. Wie wir gerade gesehen haben, können wir durchaus auch die Absenz eines Gutes als ein Übel auffassen. An diesem Punkt bietet es sich an, dass wir den Schritt vom hedonistischen Utilitarismus zum *Präferenzutilitarismus* vollziehen. Dieser scheint geradezu prädestiniert dazu, negativen Formulierungen Raum zu geben: Sein Grundbegriff kann mit ebenso gutem Recht in dem Übel einer unerfüllten Präferenz wie im Gut einer erfüllten Präferenz gesehen werden – vielleicht sogar mit größerem Recht, denn es ist keineswegs selbstverständlich, dass wir die Existenz zusätzlicher, erfüllter Präferenzen als ein Gut ansehen müssen. Wenn all meine Präferenzen erfüllt sind und ich auch keine unerfüllte Präferenz für neue Präferenzen habe – sollte ich dann das Auftreten einer neuen, wenn auch erfüllbaren, Präferenz begrüßen? Wir brauchen diese Frage hier nicht zu beantworten, sondern wollen zuerst die Plausibilität eines negativen Präferenzutilitarismus an sich prüfen, bevor wir seine möglichen Vorteile gegenüber einem positiven Utilitarismus betrachten.

Ist eine unerfüllte Präferenz immer ein Übel? Angenommen, es handle sich um eine informierte, rationale Präferenz (d. h. es liegt keine Täuschung über den subjektiven Wert der verglichenen Zustände vor), scheint diese Frage vom Standpunkt des unparteiischen Wohlwollens aus eindeutig bejahbar zu sein. Es wäre für ein Wesen offensichtlich am besten, wenn alle seine Präferenzen erfüllt würden; jede unerfüllte Präferenz schmälert seine Lebensqualität. Damit scheint sich die negative utilitaristische Forderung zu ergeben: "Minimiere unerfüllte Präferenzen!" – genauer, das Produkt aus Zahl und

durchschnittlicher Intensität unerfüllter Präferenzen.

Eine solche Position werde ich mit Christoph Fehige, einem ihrer eloquentesten Vertreter, als *Antifrustrationismus* bezeichnen (Fehige, 1998, S. 518). Wie geht sie mit den Problemen des negativen Utilitarismus um, die wir oben kennen gelernt haben? Der Antifrustrationismus lässt es zu, dass unser Leid durch eine ausreichende Menge Glück aufgewogen werden kann – dann nämlich, wenn wir die Verbindung von beidem der Absenz von beidem vorziehen würden. Dies ist sicherlich plausibler als die Haltung des negativen Utilitarismus. Er lässt es aber auch zu, dass unser Leid durch eine ausreichende Menge fremden Glückes aufgewogen werden kann, sobald die kombinierten Präferenzen der anderen unsere eigene Präferenz, unser Leid loszuwerden, an Intensität übertreffen. In der Tat läuft seine Haltung hier im Wesentlichen auf die des klassischen Utilitarismus hinaus.

Entscheidend ist aber das Verhältnis, das unsere Theorie zu den beiden Prüfsteinen Zeugung und Tod hat. Im Falle des Todes schneidet sie nun endlich befriedigend ab: Solange ein Wesen eine Präferenz für sein Weiterleben hat, gibt es einen guten Grund, es nicht zu töten – und da diese Präferenz meist sehr stark sein wird, sogar einen sehr starken Grund. Wir brauchen nicht mehr anzunehmen, wir täten jedem normalen Wesen einen Gefallen, wenn wir es schmerzlos umbrächten, wie dies im negativen Utilitarismus der Fall war, und es ist auch nicht mehr wie im negativen Utilitarismus mit Kompensation moralisch geboten, notfalls das Leben beliebig vieler glücklicher Wesen zu opfern, um einem Wesen, dessen Leben unkompensiertes Leid enthält, sein Los ein wenig zu erleichtern – die Präferenzen der anderen Wesen für ihr Weiterleben dürften normalerweise bedeutend stärker sein, und werden mit Recht als wichtiger betrachtet.

Es ist freilich denkbar, dass ein Wesen keine besondere Präferenz für sein Weiterleben besitzt (also auch keine Präferenzen, deren Erfüllung sein Weiterleben voraussetzen würde), weder als tatsächlich empfundenen Wunsch noch als hypothetische Reaktion im Falle vollständiger Informiertheit. In diesem Fall würde es der Antifrustrationismus nicht nur erlauben, sondern sogar gebieten, dieses Wesen zu töten, sofern sein zukünftiges Leben mindestens eine unerfüllte Präferenz enthalten würde – was sehr wahrscheinlich ist. Hier erweist sich unsere Theorie als Version des negativen Utilitarismus und weicht vom klassischen Utilitarismus ab. Ist ihre Haltung zu dieser Frage annehmbar? Es ist nachvollziehbar, dass die Tötung in einem solchen Fall (dessen psychologische Plausibilität fraglich ist) erlaubt sein soll; wenn das Wesen tatsächlich seinem Weiterleben keinerlei Wert beimisst, scheinen wir ihm kein Unrecht zu tun, wenn wir es töten, und es scheint vom Standpunkt des Wohlwollens aus richtig, eher ein solches Wesen zu töten als einem anderen Wesen die Erfüllung einer noch so geringen Präferenz zu versagen. Aber sollte die Tötung tatsächlich geboten sein, auch dann, wenn das Wesen zwar keine Präferenz *für*, jedoch auch keine *gegen* sein Weiterleben besitzt? Der Antifrustrationist kann antworten, dass die zukünftigen unerfüllten Präferenzen dem Wesen, wenn es denn vollständig informiert und rational wäre, einen Grund geben würden, sich den Tod zu wünschen, und wenn dieser Grund nicht (wie es normalerweise geschieht) durch eine Präferenz für das Weiterleben aufgewogen wird, ergibt sich automatisch eine Präferenz für den Tod.

Die Plausibilität dieser Antwort mag schwer zu beurteilen sein, da das Fehlen einer Präferenz entweder für oder gegen das Weiterleben ein etwas merkwürdiger und kaum alltäglicher Fall sein dürfte. Aber ein analoges Problem ergibt sich auch in einem viel

häufigeren Zusammenhang: der Frage nach der Zeugung. Hier gilt offenbar, dass wir einem ungezeugten Kind, das ungezeugt bleibt, keine Präferenz für oder gegen seine Zeugung zuschreiben können: Das Kind existiert nicht, und wenn ein Ehepaar auf Kinder verzichtet, wäre es seltsam, ihnen vorzuwerfen, sie hätten die Präferenzen ihrer nichtexistierenden Kinder unerfüllt gelassen. Wenn jedoch ein Kind gezeugt und geboren wird, wird es mit höchster Wahrscheinlichkeit erleben müssen, dass einige seiner Präferenzen nicht erfüllt werden. Das Prinzip "Minimiere unerfüllte Präferenzen!" schreibt uns also ganz klar vor, auf die Zeugung von Kindern zu verzichten, wenn diese vermutlich einige unerfüllte Präferenzen haben werden - und in der Welt, in der wir leben, dürfte dies ausnahmslos der Fall sein.

Der Antifrustrationist wird einwenden, dass ein solches Gebot in unserer Welt keineswegs mit Notwendigkeit folgen würde, da die bereits existierenden Menschen starke Präferenzen für die Zeugung von Kindern haben und ein Verzicht darauf ihre Präferenzen frustrieren würde; es könnte also das kleinere Übel sein, Kinder in die Welt zu setzen. Wie dieser Präferenzvergleich in der Praxis ausfallen würde, will ich hier nicht beurteilen; entscheidend ist, dass wir gemäß dem Antifrustrationisten aus Gründen des Wohlwollens *den Kindern gegenüber* auf ihre Zeugung verzichten sollten – alle gegenteiligen Gründe betreffen nur unser Wohlwollen gegenüber anderen Wesen. Das beste, was wir unseren Kindern tun könnten, bestünde darin, sie gar nicht erst zu zeugen. Dies läuft auf keine sehr attraktive Theorie des Wohlwollens hinaus und erscheint vielen Philosophen als fatale Schwachstelle des Antifrustrationismus (z. B. Lockwood, 1979, S. 164; Singer, 1998, S. 390f.). Es mag nicht *offensichtlich* falsch sein: Die Position, das Leben sei immer ein Übel, hat eine ebenso lange Tradition hinter sich wie die im negativen Utilitarismus mit Kompensation ausgedrückte Haltung, der Tod sei kein großes Übel; sie erscheint aber nicht konsensfähiger als diese.

Gibt es alternative Verständnisse des negativen Präferenzutilitarismus, die auch diese Klippe umschiffen können? Ich meine, ja. Der Schlüssel liegt in der Erkenntnis, dass der negative Utilitarismus, so wie er uns hier interessiert, eine mögliche Form des Prinzips des Wohlwollens ist, also eines *praktischen* Prinzips – und nicht in erster Linie ein Prinzip zur Beurteilung des Wertes verschiedener Zustände. Popper spricht davon, wie wir Leid und Glück *vom moralischen Standpunkt aus* behandeln sollen; es geht ihm nicht darum, welche Zustände Leid und welche Glück bedeuten. Es ist die Frage nach dem *moralischen* Übel – dem, was ein moralisch Handelnder zu vermeiden hat – und nicht nach dem, was an sich von Übel ist, die wir ins Zentrum stellen sollten.

Ist also, um unsere alte Frage aufzugreifen, die Existenz einer unerfüllten Präferenz in jedem Fall ein *moralisches* Übel? Wenn wir eine Präferenz ohne weiteres erfüllen könnten, dies aber nicht tun, haben wir offenbar von der Warte des unparteiischen Wohlwollens aus moralisch bedenklich gehandelt; die Existenz einer für uns erfüllbaren Präferenz scheint eindeutig ein moralisches Übel zu sein. Aber wie sieht es mit der Existenz für uns unerfüllbarer Präferenzen aus?

Eine unerfüllbare Präferenz, an deren Existenz wir nichts ändern können, ist offensichtlich kein moralisches Übel in unserem Sinne, auch wenn sie an sich ein Übel darstellen mag. Interessant sind aber die Fälle, in denen wir eine Präferenz zwar nicht erfüllen, dafür aber ihre Entstehung von vornherein verhindern können. Der Antifrustrationist wird antworten, dass wir es hier mit einem moralischen Übel zu tun haben; wenn wir ein Übel vermeiden können und dies nicht tun, verhalten wir uns

moralisch falsch. Mindestens ebenso plausibel dürfte aber die Antwort sein, es hänge ganz von dem betroffenen Wesen ab: Wenn dieses (aus welchen Gründen auch immer) eine informierte, rationale Präferenz für die Existenz der unerfüllbaren Präferenz habe, dann sei es richtig, ihr Entstehen zuzulassen, wenn es eine Präferenz dagegen habe, sei es falsch – und wenn es keinerlei diesbezügliche Präferenz habe, dann sei es eben moralisch neutral.

Die Existenz einer unerfüllbaren Präferenz kann also nach dieser Ansicht nur dann ein moralisches Übel sein, wenn ein Wesen eine Präferenz für ihre Nichtexistenz hat – und diese höherstufige Präferenz erfüllbar ist! Somit wäre nach unserer neuen Theorie im Bereich des Wohlwollens jedes moralische Übel auf die *Nichterfüllung einer erfüllbaren Präferenz* zurückzuführen, während für den Antifrustrationisten die *Nichtvermeidung einer vermeidbaren unerfüllten Präferenz* die Wurzel allen moralischen Übels ist. "Minimiere unerfüllte erfüllbare Präferenzen!" – dies könnte unser neuer, zugegebenermaßen nicht mehr sehr griffiger Slogan sein, "Wähle stets diejenige Alternative, bei der das Produkt aus Zahl und durchschnittlicher Intensität unerfüllter Präferenzen, die du durch eine andere Wahl erfüllen könntest, am geringsten ist!" seine Ausdeutung.

Da es schwer sein dürfte, auf einer rein theoretischen Ebene zwischen diesen beiden Alternativen zu entscheiden, wollen wir uns erneut anschauen, wie unsere neue Theorie mit den bekannten Klippen für den negativen Utilitarismus fertig wird. In den meisten Fällen wird ihre Haltung nicht von derjenigen des Antifrustrationismus abweichen, so dass wir uns auf die Fälle konzentrieren können, in denen dieser in Schwierigkeiten gerät. Die Tötung eines Wesens, das keinerlei Präferenz für oder gegen sein Weiterleben besitzt, wird unsere Theorie weiterhin erlauben, aber nicht mehr fordern: Wenn wir das Wesen töten, haben wir keine für uns erfüllbaren Präferenzen unerfüllt gelassen; es ist zwar möglich, dass das Wesen in Zukunft Präferenzen gehabt hätte, die für uns erfüllbar gewesen wären, aber so wie die Dinge liefen, waren diese Präferenzen zu keinem Zeitpunkt existent (weder als tatsächlicher Wunsch noch als hypothetische Reaktion). Lassen wir das Wesen hingegen am Leben, haben wir ebenfalls keine für uns erfüllbaren Präferenzen unerfüllt gelassen; es wird zwar wohl mit einigen unerfüllten Präferenzen weiterleben, aber diese hätten wir nicht erfüllen können, und wir haben auch keine erfüllbare Präferenz dafür, diese Präferenzen nicht zu haben, unerfüllt gelassen, da das Wesen eine solche Präferenz (die *de facto* eine Präferenz gegen sein Weiterleben gewesen wäre) nicht hatte.

Analog ist unser Ergebnis bei der deutlich wichtigeren Frage nach der Beurteilung der Zeugung. Wie für den Antifrustrationismus ist für unsere Theorie der Verzicht auf die Zeugung eines Kindes kein moralisches Übel (zumindest nicht aus Gründen, die das Wohlwollen dem Kind gegenüber betreffen): Die erfüllbaren Präferenzen, die das Kind im Falle seiner Existenz gehabt hätte, werden nie real, und so kann man den Eltern auch nicht vorwerfen, sie nicht erfüllt zu haben. Andererseits ist aber auch die Zeugung eines Kindes kein moralisches Übel mehr: Zwar wird das Kind in seinem Leben voraussichtlich zahlreiche unerfüllte Präferenzen haben, aber diese hätten wir unmöglich erfüllen können. Im Normalfall wird es uns also freigestellt sein (oder lediglich von Überlegungen zum Wohl der bereits existierenden Bevölkerung abhängen), ob wir Kinder zeugen wollen oder nicht. Dies ist sicher ein ungemein plausibles Ergebnis.

Zu präzisieren wäre freilich, dass es nach unserer Theorie durchaus Situationen gibt, in denen das Zeugen von Kindern aus Gründen des Wohlwollens ihnen selbst gegenüber moralisch unzulässig ist. Dies ist insbesondere dann der Fall, wenn im Leben eines Kindes

voraussichtlich das Leid das Glück überwiegen würde – wir also voraussehen können, dass sein Leben unkompensiertes Leid enthalten wird. In die Terminologie des Präferenzutilitarismus übertragen hieße dies, dass das Kind eine Präferenz haben würde, nicht gezeugt zu werden. Wenn wir es nun dennoch zeugten, wäre diese durch die Zeugung des Kindes real gewordene Präferenz unerfüllt geblieben, obwohl wir sie leicht hätten erfüllen können. Auch diese Konsequenz unserer Theorie dürfte intuitiv annehmbar sein.

Es gibt allerdings auch etwas kompliziertere Situationen, in denen unsere Theorie das Zeugen von Kindern als problematisch erscheinen lässt: Wenn etwa Eltern, um alle für sie erfüllbaren Präferenzen eines Kindes zu erfüllen, ihre eigenen Präferenzen in einem solchen Ausmaß unerfüllt lassen müssten, dass sie es vorziehen würden, kein Kind zu haben, dann wäre es ihnen nach unserer Theorie nicht gestattet, das Kind zu zeugen im Wissen, dass sie um ihres eigenen Vorteils willen einige seiner Präferenzen unerfüllt lassen würden – es sei denn, ihre eigene Präferenz, ein Kind zu haben, wäre stärker als die zukünftigen unerfüllt bleibenden erfüllbaren Präferenzen des Kindes. Verdeutlichen wir uns dies erneut mit Zahlenwerten, in drei verschiedenen Beispielen:

	w_1	w_2	w_3
e_1	-1	0	-2
k_1	-	-3	0
e_2	-2	0	-3
k_2	-	-1	0
e_3	-1	0	-3
k_3	-	-2	0

Mit w_1 bezeichne ich jeweils die mögliche Welt, in der das Kind nicht gezeugt wird, w_2 ist die für die Eltern optimale Welt, in der sie nicht zugunsten des Kindes auf die Erfüllung wichtiger eigener Präferenzen verzichten, und w_3 ist die für das Kind optimale Welt, in der all seine für die Eltern erfüllbaren Präferenzen erfüllt werden; der Einfachheit halber nehme ich an, dass dies alle realisierbaren Möglichkeiten sind. e_1 bis e_3 sind jeweils die Eltern, k_1 bis k_3 ist jeweils das Kind, und die Zahlen stehen für das Produkt aus Zahl und durchschnittlicher Intensität unerfüllter erfüllbarer Präferenzen (der Anschaulichkeit halber als negative Zahlen dargestellt). Ich nehme jeweils an, dass die Eltern aufgrund ihrer Präferenzen von selbst nie w_3 statt w_1 wählen würden, und offensichtlich dürften wir sie auch nicht dazu zwingen (da in allen drei Fällen das Ausmaß der Präferenzfrustration höher als in w_1 wäre). Es bleibt also, w_1 und w_2 miteinander zu vergleichen.

Im ersten Fall (der eine Version von Derek Parfits *Mere Addition Paradox* darstellt; cf. Parfit, 1987, S. 419-441) wäre es den Eltern nicht gestattet, ein Kind zu haben, und dies dürfte intuitiv plausibel sein: Es ist unter den gegebenen Umständen den Eltern eher zuzumuten, auf ihren Kinderwunsch zu verzichten, als dem Kind, mit einem hohen Maß an eigentlich erfüllbaren, aber unerfüllten Präferenzen auf die Welt zu kommen. Wenn wir w_2 mit w_3 vergleichen, ist das Kind in w_2 eindeutig einem Unrecht ausgesetzt; seine Lage hätte in einem größeren Ausmaß besser sein können, als dies das Schicksal anderer verschlechtert hätte. Ich denke, dass uns unsere Theorie hier die richtige Antwort gibt.

Der zweite Fall dürfte ziemlich unproblematisch sein: Hier tritt die oben angesprochene Möglichkeit ein, dass der Kinderwunsch der Eltern stark genug ist, um die Wahl von w_2 zu rechtfertigen. Das Kind muss sich zwar mit einigen unerfüllten erfüllbaren Präferenzen abfinden, aber diese wären nur unter unverhältnismäßig hohen Kosten für die Eltern erfüllbar gewesen. Bei entsprechend starkem Kinderwunsch muss es Eltern gestattet sein, auch dann ein Kind zu

haben, wenn sie nicht schwerwiegende eigene Interessen zugunsten relativ unbedeutender Interessen des Kindes zu opfern bereit sind. Erneut ein bestandener Test für unsere Theorie.

Interessant ist aber der dritte Fall, der Züge der ersten beiden kombiniert: Zum einen würde wie im ersten Fall der Verzicht auf ein Kind die Präferenzen der Eltern weniger frustrieren als die Wahl von w_2 diejenigen des Kindes. Zum anderen aber wiegen bei einem Vergleich von w_2 und w_3 wie im zweiten Fall die Präferenzen der Eltern schwerer; sollte das Kind gezeugt werden, so hätte es kein Recht, auf der Erfüllung seiner Präferenzen zu bestehen. Unsere Theorie sagt hier, dass die Zeugung des Kindes unter diesen Umständen nicht erlaubt ist: Dass das Kind unerfüllte erfüllbare Präferenzen haben wird, bleibt ein Übel; ein solches kann zwar wie im zweiten Fall gelegentlich durch einen starken Kinderwunsch gerechtfertigt werden, aber es kann auch Fälle geben wie den gegenwärtigen, in denen die Präferenzen der Eltern zu schwach dafür sind.

Ich halte diese Antwort für vertretbar, aber auch die gegenteilige Antwort hat intuitiv eine gewisse Anziehungskraft, und wenn wir sie verteidigen wollten, würden wir uns am besten auf folgende Überlegung berufen: Die Präferenzen des Kindes wären nur erfüllbar, wenn dafür diejenigen anderer Wesen in stärkerem Maße vernachlässigt würden; das heißt, sie wären nur durch eine *moralisch unzulässige* Handlung erfüllbar, weshalb ihre Nichterfüllung kein moralisches Übel ist. Dies führt uns zu einer noch weiter modifizierten Version des negativen Utilitarismus, die sich auf eine These von T. M. Scanlon stützen kann. Nach Scanlon ist eine Handlung genau dann moralisch falsch, wenn ein betroffenes Wesen die entsprechende Handlungsregel mit guten Gründen als moralisches Prinzip zurückweisen könnte. Diese guten Gründe können etwa gegeben sein, wenn eine Handlungsregel mich unnötigerweise mehr belastet als andere; sie sind aber ausdrücklich nicht gegeben, wenn die Handlungsregel mir zwar eine Last auferlegt, jede für mich bessere Alternative aber andere Wesen noch mehr belasten würde (Scanlon, 1982, S. 110f.). Eine solche Theorie würde sich insofern vom Utilitarismus entfernen, als sie nicht mehr durch ein Maximierungs- oder Minimierungsgebot ausgedrückt wird; es wäre nicht mehr zu fragen, bei welcher Handlung die Nichterfüllung von Präferenzen minimiert würde, sondern bei welcher Handlung *kein* Wesen Einwände erheben könnte. Wenn wir jedoch (was Scanlon selbst nicht tun würde) die möglichen Einwände auf das Aufzeigen von Verstößen gegen die Minimierung unerfüllter erfüllbarer Präferenzen beschränken würden, wäre der Utilitarismus durch die Hintertür wieder eingeführt, und wir könnten weiterhin von einer Version des negativen Utilitarismus sprechen.

Wenn wir uns nun der Situation des Kindes in unserem dritten Fall zuwenden, erkennen wir, dass das Kind bei einer Wahl von w_2 keine guten Gründe für eine Beschwerde hätte: Es kann sich nicht darüber beschweren, dass statt w_2 nicht w_3 realisiert wurde, da in w_3 das Ausmaß unerfüllter erfüllbarer Präferenzen höher wäre; und solange es sein Leben in w_2 immer noch für erlebenswert hält, *wird* es sich nicht darüber beschweren, dass nicht w_1 realisiert wurde. Würde freilich das Leben des Kindes durch die Nichterfüllung seiner Präferenzen in w_2 völlig unerträglich, dann würde es mit guten Gründen für die Wahl von w_1 plädieren können, und die Wahl von w_2 durch die Eltern wäre unzulässig.

Es sei betont, dass sich diese neue Version des negativen Präferenzutilitarismus in den meisten anderen Situationen nicht von der davor behandelten unterscheidet; insbesondere kommen beide Theorien in all jenen Fällen (und das sind die meisten) zum gleichen Resultat, in denen es nicht um die Erzeugung zusätzlicher Wesen, sondern nur um Verteilungsprobleme zwischen bereits

existierenden Wesen geht. Ich kann die relativen Verdienste der beiden hier nur knapp skizzierten Theorien in diesem Aufsatz nicht weiter erörtern (was auch eine eingehende Diskussion der ihnen jeweils zugrunde liegenden Moralverständnisse erfordern würde); ich glaube aber, dass beide eine nähere Ausarbeitung und Prüfung lohnen würden.

*

Wir sind letztlich bei zwei Theorien angelangt, die sich von der ursprünglichen Form des negativen Utilitarismus stark unterscheiden, sich aber doch auf seine Ideen berufen: Sie gehen von der Vermeidung von Übel, nicht der Vermehrung von Gutem als Grundprinzip der Moral (oder zumindest der Theorie des Wohlwollens) aus, und sie sind eindeutig utilitaristisch, was Natur und Aggregation des Guten bzw. des Übels betrifft. Stellen diese Formen des negativen Utilitarismus nun einen echten Fortschritt gegenüber dem klassischen Utilitarismus dar, und bilden sie eine brauchbare Grundlage für eine Theorie des Wohlwollens?

Wie wir gesehen haben, fallen die Antworten unserer Theorien in den meisten Fällen mit denjenigen des gewöhnlichen Utilitarismus zusammen, insbesondere dort, wo es um einfache Verteilungsprobleme geht. Es wird weiterhin erlaubt sein, einem Wesen Leid zuzufügen, wenn dadurch einem anderen eine größere Menge Glück beschert wird; einem der Hauptkritikpunkte am klassischen Utilitarismus, der Poppers Überlegungen zugrunde lag, wird damit nicht Rechnung getragen. Es könnte dem Utilitaristen jedoch leichter fallen, sich auf unsere Theorien stützend gegen diese Kritik zu verteidigen. Zum einen macht die Konzentration auf Präferenzen die Zurückweisung einer klaren Grenze zwischen Glück und Leid plausibler: Wenn wir es gelegentlich tatsächlich vorziehen, Glück zu erlangen, statt von Leid frei zu bleiben, dann sollte unsere Theorie des Wohlwollens dies auch widerspiegeln. Vor allem aber streicht die negative Formulierung heraus, dass es bei der Förderung von Glück auch um die Vermeidung eines Übels, nämlich der Existenz unerfüllter erfüllbarer Präferenzen, geht. Wenn wir in einigen Fällen statt der Vermeidung geringen Leides das Glück anderer in hohem Ausmaß fördern, lassen wir nicht ein Übel zugunsten eines Gutes außer Acht, sondern entscheiden uns unter zwei Übeln für das geringere, das wie gesagt unseren gewöhnlichen Präferenzen nach zu urteilen gelegentlich das Zufügen oder Zulassen eines Leides sein kann.

Der Bereich, in dem sich unsere Theorien stark vom klassischen, positiven Utilitarismus unterscheiden, ist derjenige der Erzeugung zusätzlicher Wesen. Bei diesem handelt es sich nun um ein stark vermintes Gelände, auf welchem der klassische Utilitarismus ebenso wie die meisten anderen Theorien des Wohlwollens auf größte Schwierigkeiten stößt (cf. vor allem Parfit, 1987, S. 351-454); für einen positiven Utilitaristen, der Glück oder Präferenzerfüllung als ein Gut an sich ansieht, droht immer die unangenehme Konsequenz, dass das Zeugen von möglichst vielen Kindern zur moralischen Pflicht wird. Was wir von unseren Theorien in diesem Bereich gesehen haben, lässt dagegen zuversichtlich sein, so dass der negative Präferenzutilitarismus hier die dringend benötigte Kur für den Utilitarismus darstellen könnte.

Ich komme also aus diesen beiden Gründen zum Schluss, dass ein Utilitarist tatsächlich ein negativer Utilitarist sein sollte, im Sinne einer unserer beiden zuletzt entwickelten Theorien oder einer weiteren Abwandlung davon. Inwieweit unser negativer Präferenzutilitarismus auch für Philosophen, die von der Commonsense-Ethik herkommen, eine Alternative darstellt, bleibt hingegen fraglich: Die Hoffnungen auf eine in ihren Forderungen "bescheidenere" und die "Opferung" Unglücklicher vermeidende Version des Utilitarismus, die der negative

Utilitarismus ursprünglich zum Ausdruck brachte, haben sich nicht erfüllt. Aber die Stärkung, die der Utilitarismus durch den Übergang zur negativen Version erfährt, sollte auch für die Gegner dieser Theorie ein neue Herausforderung darstellen.

Literatur:

FEHIGE, Christoph: "A Pareto Principle for Possible People", in: Christoph FEHIGE und Ulla WESSELS (Hrsg.), *Preferences*, Berlin / New York 1998.

GRIFFIN, James: *Well-Being*, Oxford 1986.

LOCKWOOD, Michael: "Singer on Killing and the Preference for Life," *Inquiry* 22 (1979), 157-170.

PARFIT, Derek: *Reasons and Persons*. Reprinted with further corrections, Oxford ^{3}1987.

POPPER, Karl R.: *The Open Society and Its Enemies*, London 1945.

SCANLON, T. M.: "Contractualism and utilitarianism", in: Amartya SEN und Bernard WILLIAMS (Hrsg.), *Utilitarianism and beyond*, Cambridge 1982.

SCARRE, Geoffrey: *Utilitarianism*, London 1996.

SHAW, William H.: *Contemporary Ethics: Taking Account of Utilitarianism*, Oxford 1999.

SINGER, Peter: "Possible Preferences", in: Christoph FEHIGE und Ulla WESSELS (Hrsg.), *Preferences*, Berlin / New York 1998.

SMART, J. J. C.: *An outline of a system of utilitarian ethics*, in: J. J. C. SMART und Bernard WILLIAMS, *Utilitarianism: For and Against*, Cambridge 1973.

SMART, R. N.: "Negative Utilitarianism", *Mind* 67 (1958), 542-543.

WALKER, A. D. M.: "Negative Utilitarianism", *Mind* 83 (1974), 424-428.