

## Parfit und Kant über vernünftige Zustimmung<sup>1</sup>

MARTIN STICKER, GÖTTINGEN / TRINITY COLLEGE DUBLIN

*Zusammenfassung:* Nach Parfit konvergieren die systematisch stärksten Versionen von Kantianismus, Regel-Konsequentialismus und Kontraktualismus in einer Triple Theory. Ich konzentriere mich auf eine der zentralen Schwierigkeiten, Kantianismus und Konsequentialismus zusammenzubringen: die Rolle von Zustimmung, welche ihren deutlichsten Ausdruck in Kants Zweck-an-sich-Formel findet. Ich zeige zunächst, wie die Einführung unparteilicher, nichtmoralischer Gründe, auf der viel Gewicht in Parfits Zustimmungsprinzip liegt, in einigen Fällen die Zweck-an-sich-Formel zu dem intuitiv richtigen Ergebnis führen kann. Anschließend wende ich mich kritisch gegen Parfit. Ich diskutiere zwei Einwände gegen Parfits Zustimmungsprinzip: Das Zustimmungsprinzip ist unterbestimmt und daher nicht relevant für die Fälle, die Kantianismus und Konsequentialismus unterscheiden. Zudem lässt das Zustimmungsprinzip, wie Parfit es versteht, mehr Instrumentalisierung zu, als Parfit selbst bereit ist einzuräumen.

*Schlagwörter:* Triple Theory, Zustimmung, Kantianismus, Konsequentialismus, Derek Parfit

---

1 Ich danke Matthias Hoesch und den anonymen Gutachtern der Zeitschrift für praktische Philosophie für zahlreiche Anmerkungen und Einwände zu einer früheren Version dieses Artikels. Ich habe eine frühere Version dieses Artikels auf der Konferenz: „Parfits *On What Matters* in der Diskussion“ an der Universität Frankfurt vorgestellt. Ich danke den Veranstaltern Philipp Schink und Achim Vesper sowie den Teilnehmern für Hinweise, Einwände und Diskussion.

In der Einleitung von *On What Matters* (OWM<sub>1</sub>, xxxiii) bezeichnet Parfit Kant als den wichtigsten Moralphilosophen seit den alten Griechen. Zudem behauptet er später, dass „in the cascading fireworks of a mere forty pages [der *Grundlegung*, MS], Kant gives us more new and fruitful ideas than all the philosophers of several centuries. Of the qualities that enable Kant to achieve so much, one is inconsistency“ (OWM<sub>1</sub>, 183). In *On What Matters* will Parfit aufzeigen, dass Kants Ethik, wenn man von dem absieht, was an ihr inkonsistent oder schlichtweg falsch ist, mit zwei anderen bedeutenden Ansätzen der Ethik, Konsequentialismus und Kontraktualismus, vereinbar sei. Im vorliegenden Aufsatz möchte ich einen Aspekt von Parfits Versuch, Kantianismus und Konsequentialismus zu einem einheitlichen Prinzip bzw. einer neuen Theorie zu entwickeln, kritisch diskutieren: Die Idee, dass wir vernünftige Akteure (andere wie auch uns selbst) nie nur als Mittel, sondern immer auch als Zweck zu behandeln haben. Diese Idee scheint in gegenwärtigen Debatten für viele Moralphilosophen der wichtigste Gedanke der kantischen Ethik zu sein,<sup>2</sup> und dementsprechend ist Parfits kritische Diskussion dieser Idee zentral für die Bewertung Parfits aus einer kantischen Perspektive.

Parfits ausführliche Diskussion der Zweck-an-sich-Formel beansprucht den Großteil von Band I, Teil 2 von OWM. Ich kann diese Diskussion nicht in ihrer Gänze behandeln und konzentriere mich auf Parfits Interpretation des *Zustimmungsprinzips*, welches verlangt, dass wir andere nicht auf eine Art und Weise behandeln, der sie nicht vernünftigerweise zustimmen können. Parfit selbst betont: „It is surprising that this

---

2 Siehe etwa Wood (1998, 165), der behauptet, dass die Zweck-an-sich-Formel „the greatest resonance with our culture’s moral consciousness“ habe. Parfit erklärt, dass das „requirement to respect all persons is one of Kant’s greatest contributions to our moral thinking“ (OWM<sub>1</sub>, 233).

principle has been so little discussed. The principle has great appeal, and is worth considering as a separate moral idea, not merely as another way of stating Kant's Formula of Universal Law" (OWM1, 183). Vorliegender Aufsatz diskutiert die Bedeutung dieses Prinzips im Kontext von Parfits Versuch, wesentliche kantische Ideen in seine Theorie zu integrieren. An meiner Diskussion werden einige der zentralen Schwierigkeiten deutlich werden, Ethiken, welche Pflicht von der Autonomie und der Selbstzweckhaftigkeit des Akteurs her denken, mit konsequenzbasierten ethischen Ansätzen zu vereinigen.

In einem einführenden Teil möchte ich kurz erklären, was das Ziel der Diskussion von Kantianismus und Konsequentialismus in *On What Matters* ist sowie welche Funktion die Diskussion der Zweck-an-sich-Formel im Kontext von Parfits Projekt erfüllt. In einem zweiten Teil erläutere ich zunächst, wie Parfits Zustimmungsprinzip und seine Theorie unparteilicher, nicht-moralischer Gründe in einigen Fällen zeigen können, wie die Zweck-an-sich-Formel zu dem intuitiv richtigen Ergebnis führt. Dann erhebe ich zwei Einwände gegen Parfits Zustimmungsprinzip: Das Zustimmungsprinzip ist *unterbestimmt* und daher nicht relevant für die Fälle, die Kantianismus und Konsequentialismus unterscheiden. Zudem lässt das Zustimmungsprinzip, wie Parfit es versteht, *mehr Instrumentalisierung* zu, als Parfit selbst bereit ist einzuräumen. Das Ergebnis meiner Diskussion fällt kritisch aus gegenüber Parfits Versuch, Kantianismus und Konsequentialismus einander anzulehnen.

## 1.

Dieser einführende Abschnitt hat drei Ziele. Erstens stelle ich Parfits Theorie unparteiischer Gründe vor, da diese Form der Gründe von entscheidender Bedeutung für die Anwendung

des Zustimmungsprinzips ist. Zweitens werde ich die Rolle von tiefen moralischen Meinungsverschiedenheiten (*deep moral disagreements*) erläutern, da es eines der Hauptanliegen von Parfits Diskussion der Zweck-an-sich-Formel ist, solche Meinungsverschiedenheiten auszuräumen bzw. als lediglich scheinbare Meinungsverschiedenheiten zu entlarven. Ob ihm dies gelingt, stellt einen Prüfstein für Parfits Diskussion dar. Drittens werde ich einige einführende und kontextualisierende Bemerkungen zur Debatte um den kantischen Konsequentialismus hinzufügen und herausarbeiten, mit welchen Problemen sich Parfits Triple Theory konfrontiert sieht. Dies wird zu einer kritischen Diskussion des Parfit'schen Zustimmungsprinzips überleiten.

*On What Matters* stellt Parfits Versuch dar, eine objektivistische Theorie von Gründen zu etablieren und gegen subjektivistische Theorien zu verteidigen. Im Zuge seiner Diskussion von Gründen in OWM1 stellt Parfit drei Theorien vor, diese sind: (i) Rationaler Egoismus (*Rational Egoism*), eine Position, die behauptet, dass ich überwiegende Gründe habe zu tun, was am besten für mich selbst ist (was am besten meine Präferenzen oder Wünsche erfüllt); (ii) Vernünftige Unparteilichkeit (*Rational Impartialism*), eine Position, die behauptet, dass ich überwiegende Gründe habe zu tun, was unparteilich betrachtet das Beste ist; und (iii) den *Dualismus der praktischen Gründe*. Letztere Position geht auf Henry Sidgwick zurück und ist nach Parfit der Wahrheit am nächsten (OWM1, 131). Laut Parfits Weiterentwicklung dieses Dualismus, welcher in seiner Diskussion des Zustimmungsprinzips wichtig werden wird, habe ich oft hinreichende Gründe, sowohl das zu tun, was für mich persönlich das Beste ist, als auch das, was unparteilicherweise das Beste ist.

Interessanterweise glaubt Parfit, dass Vernünftige Unparteilichkeit in der Form von unparteiisch-vernünftigem Handlungskonsequentialismus (Impartial-Reason Act Consequentialism) ein externer Rivale der Moral sei („external rival to morality“, OWM1, 168) und nicht bereits eine moralische Position darstelle. Dies ist der Fall, da es dem unparteiisch-vernünftigen Handlungskonsequentialismus nicht darum geht, was *Pflicht* ist, sondern, wozu wir unparteiliche Gründe haben, d.h., es geht ihm um eine bestimmte Form der Gründe, nicht um Pflicht oder Moral. Parfit kontrastiert Vernünftige Unparteilichkeit sogar mit Moralischem Rationalismus (*Moral Rationalism*), einer Position, die behauptet, dass ich überwiegende Gründe habe, meine Pflicht zu tun. Es ist sehr wichtig für das Verständnis von Parfits Position, dass die Beförderung des unparteilichen Guten (Vernünftige Unparteilichkeit) und der Pflichterfüllung (Moralischer Rationalismus) nicht zusammenfallen. Ein Handlungskonsequentialist wie Sidgwick kann beispielweise unparteilich sein, wenn er gemäß seiner handlungskonsequentialistischen Überzeugungen handelt, muss aber nicht unbedingt ein moralischer Rationalist sein, da Handlungen, welche das unparteilich Gute befördern, moralisch verboten sein können.<sup>3</sup> Wir werden auf die Spannung zwischen unparteilicher Gutheit und Moral noch ausführlich zu sprechen kommen.

---

3 Auf Seite 136 von OWM1 sagt Parfit klar, dass Handlungen falsch sein können, sogar wenn sie das unparteiliche Gute maximal befördern. Parfit will zeigen, dass es nicht darum geht, dass „ultimately matters [...] how well things go“ (OWM1, 26), da dies oft kontraintuitive Implikationen hätte. Siehe auch: „There are many Utilitarian acts to which some people could not rationally consent, and many non-Utilitarian acts to which everyone could rationally consent“ (OWM1, 191) sowie OWM1, Abschn. 63.

Als Teil seiner Ausarbeitung einer objektivistischen Theorie von Gründen will Parfit zeigen, dass es moralische Wahrheiten in einem robusten, nicht subjektabhängigen Sinn gibt.<sup>4</sup> Seine *Triple Theory*, welche Parfit in Buch I von OWM entwickelt, kann als Antwort auf das verstanden werden, was er als das größte Problem einer objektivistischen moralischen Theorie sieht: tiefe moralische Meinungsverschiedenheiten. Diese wären gegeben, wenn verschiedene Akteure ohne einem Missverständnis erlegen zu sein mit gleicherweise vernünftiger Berechtigung zu konfligierenden Ansichten über Pflicht oder Moral kämen. Wäre dies möglich, dann wäre dies ein starker Hinweis darauf, dass moralische Gründe von den persönlichen Präferenzen oder Einstellungen der Akteure abhingen, d.h. nur im Rahmen einer subjektivistischen Theorie richtig zu verstehen seien. Gegen einen möglichen Einwand, der sich aus tiefen moralischen Meinungsverschiedenheiten ergibt, will Parfit zeigen, dass es eine einheitliche moralische Theorie gibt, welche sich aus den systematisch stärksten Versionen von sich vermeintlich in Spannung miteinander befindenden Theorien ergibt.<sup>5</sup>

Damit sollte auch klar sein, was für Parfit in seiner Diskussion Kants auf dem Spiel steht. Parfit muss zeigen, dass zentrale Elemente des Kantianismus mit der plausibelsten Form des Konsequentialismus vereinbar sind. Andernfalls könnte es der Fall sein, dass es keine moralischen Wahrheiten *simpliciter* gibt, sondern bestenfalls Wahrheiten innerhalb bestimmter

---

4 Parfit legt sich nicht darauf fest, dass jeder, der eine objektivistische Grundtheorie vertritt, auch annimmt, dass es moralische Wahrheiten in einem nicht subjektabhängigen Sinne gibt. Diese Annahme, wenn sie begründet werden kann, stärkt aber den Objektivismus bzw. stellt ein Problem für den Subjektivismus dar.

5 Siehe OWM2, 155: „*In trying to combine these different kinds of moral theory, my main aim was not to find a supreme principle, but to find out whether we can resolve some deep disagreements*“.

Theorien. Moralische Wahrheit wäre dann möglicherweise davon abhängig, ob ein Akteur lieber Kantianer oder Konsequentialist ist bzw. ob er kantische oder konsequentialistische Intuitionen oder Überzeugungen hat. Laut Parfit gäbe es damit gar keine moralischen Wahrheiten mehr und eine Form des Nihilismus wäre die einzig angemessene Position.

Parfits Strategie besteht oftmals darin, eine Position oder ein Prinzip vorzuführen und dann mit Einwänden, wie etwa gegenläufigen Intuitionen bezüglich spezifischer Anwendungsfälle, zu konfrontieren und anhand dieser Einwände weiterzuentwickeln. Die avancierteste Version dieser Position oder dieses Prinzips schlägt schließlich in eine andere Position um oder fällt mit dieser zusammen. Parfits zeitweise quasi-dialektische Vorgehensweise macht es schwer herauszufinden, was Parfit selbst über die Ideen eines Objektivismus und des Prinzips der Triple Theory hinaus eigentlich vertritt und was er nur auf dem Weg zu diesen Resultaten annimmt und diskutiert.

Diese Schwierigkeit ist reflektiert in der Frage, welche Rolle Parfits Diskussion der Zweck-an-sich-Formel in OWM, Band I, Teil 2 eigentlich spielt. In Teil 1 stellt Parfit die wichtigsten Grundgedanken seines Objektivismus vor und entwickelt sein begriffliches Instrumentarium und in Teil 3 diskutiert er Kandidaten für oberste Moralprinzipien und führt verschiedene dieser Prinzipien auf ein einheitliches Prinzip zurück. In Teil 2 diskutiert Parfit Prinzipien oder Ideen, welche sich nicht als Kandidat für das oberste Moralprinzip eignen, welche aber dennoch für eine erfolgreiche Vereinigung der obersten Prinzipien in Teil 3 abgehandelt werden müssen. Parfit betont, dass wir gute Gründe haben, eine Form des Zustimmungsprinzips zu akzeptieren:

in most cases, it is wrong to act in ways to which anyone could not rationally consent. When our acts would affect many people, and there is only one possible act to

which everyone could rationally consent, this fact gives us a strong reason to act in this way, and may be enough to explain why such acts are morally required. And on some plausible assumptions, the Consent Principle could never go astray, by requiring acts that are wrong for other reasons, or condemning acts that are required (OWM1, 211).<sup>6</sup>

Das Zustimmungsprinzip in seiner entwickelten Form ist, laut Parfit, einer der tragfähigsten Bestandteile von Kants Theorie. Es verdient damit Diskussion als ein eigenständiges Moralprinzip. Auch wenn Parfit das Zustimmungsprinzip letztlich nicht weiterentwickelt und er sich in OWM3 auf die universale Gesetzesformel stützt, sollte die Bedeutung von Parfits Diskussion der Zweck-an-sich-Formel nicht unterschätzt werden. Das Zustimmungsprinzip ist bedeutsam für Parfit, da er sicherstellen muss, dass sich Zustimmung als Teil einer Ethik mit Prinzipien, die konsequentialistisch in einem weiten Sinne sind, lesen lässt. Parfit will in OWM2 zeigen, dass die Zweck-an-sich-Formel Kantianismus und Konsequentialismus nicht unvereinbar macht, bzw. er will ausräumen, was als das größte Problem einer Vereinigung von Kantianismus und Konsequentialismus gesehen werden kann, nämlich dass Kantianer Ethik von der Autonomie und der Würde des Subjekts her denken, Konsequentialisten hingegen von einem unparteilichen Guten, welches zu befördern ist. Dementsprechend sind Konsequentialisten tendenziell bereit, ein Individuum um der Beförderung eines unparteilichen Guten willen zu instrumentalisieren, wo-

---

6 Parfit äußert sich allerdings gelegentlich auch skeptisch bezüglich dieses Prinzips: „The Consent Principle cannot, however, be what Kant was trying to find: the supreme principle of morality. Some acts are wrong even though everyone could rationally consent to it“ (OWM1, 211).

hingegen Kantianer ein strenges Instrumentalisierungsverbot annehmen.

Für Parfits Vorhaben einer einheitlichen Theorie ist es entscheidend, ob er tiefe moralische Meinungsverschiedenheiten in Fällen, in denen plausible Versionen von Kantianismus und Konsequentialismus in ihren Urteilen auseinanderzufallen scheinen, ausräumen kann. Entsprechend Parfits Anliegen und der Funktion von Teil 2 für dieses Anliegen können wir daher sagen, dass Parfits Diskussion der Zweck-an-sich-Formel dann und nur dann ein Erfolg ist, wenn er ein Prinzip – oder eine Position – entwickelt, das die Möglichkeit von tiefen moralischen Meinungsverschiedenheiten, welche sich aus den in OWM2 diskutierten moralischen Problemfällen ergibt, plausibel verneinen kann.

Parfits Idee, dass zentrale Elemente des Kantianismus und Formen des Konsequentialismus miteinander vereinbar sind, findet sich nicht erstmals in *On What Matters*. Richard Hare (1993) argumentiert beispielsweise für die Möglichkeit eines kantischen Utilitarismus, und David Cummiskey (1990, 588) behauptet, Kant „actually provides support for a form of normative consequentialism“ (vgl. auch Cummiskey 1996). Oft ist die Grundidee einer Verbindung von Kantianismus und Konsequentialismus, dass Kants Ethik eine fruchtbare Moralphyschologie, Metaethik und Moralbegründung enthält, dass seine moralischen Prinzipien in manchen konkreten Fällen aber kontraintuitive Resultate haben und daher durch konsequenzsensitivere Prinzipien ersetzt werden sollten bzw. Handlungsfolgen auf der Ebene konkreter Verpflichtungen stärker in Kants Theorie integriert werden müssen.<sup>7</sup>

---

7 Cummiskey (1990, Abschn. 2) unterscheidet ausdrücklich zwischen einem Konsequentialismus auf Ebene der Normen und einem Konse-

Die Debatte um die Möglichkeit eines kantischen Konsequentialismus hat eine Reihe von wichtigen Unterschieden zwischen Kantianismus und Konsequentialismus aufgezeigt. Dazu gehören die Fragen, wem gegenüber wir Pflichten haben oder was der moralische Status von vernunftlosen, leidensfähigen Wesen ist, ob es Pflichten gegen sich selbst gibt und was deren Rolle ist, wie unparteilich Moral sein darf und muss, was das Objekt ethischer Evaluationen ist (Folgen, Handlungen, Regeln oder Maximen) und was die Rolle moralischer Motivation ist.<sup>8</sup> Zwei Streitpunkte sind für vorliegenden Aufsatz von besonderem Interesse:

(i) *Können Pflichten die Befolgung des Guten einschränken?* Einer der Hauptgründe, warum Kants Ethik vielen zeitgenössischen PhilosophInnen attraktiv erscheint, ist die in ihr entfaltete Idee, dass wir bestimmte Handlungen nie ausführen dürfen, ungeachtet ihrer Konsequenzen. Diese Verbote speisen sich bei Kant aus zwei Quellen: Zum einen sind für Kant vernünftige Wesen nie wie bloße Sachen zu behandeln und dürfen daher nicht instrumentalisiert werden. Zum anderen glaubt Kant, dass sogenannte „strenge Pflichten“, welche sich daraus ergeben, dass eine Maxime nicht ohne Widerspruch im Denken universalisiert werden kann, gegenüber sogenannten „weiten Pflichten“, welche sich daraus ergeben, dass eine Maxime nicht als universalisiert gewollt werden kann, Priorität genießen (vgl. IV: 424). Aus dieser Priorität folgt, dass es Dinge gibt, die wir niemals tun dürfen – etwa Versprechen brechen –, sogar wenn

---

quentialismus auf Ebene der Begründung dieser Normen. Laut Cummiskey ist Kant nur in ersterem Sinn ein Konsequentialist.

8 Siehe Hoesch, Sticker (im Erscheinen) für eine Diskussion einiger dieser Aspekte. Die meisten dieser Aspekte werden auch von Timmermann (2005) diskutiert. Er vertritt die Auffassung, dass Kants Theorie „essentially anti-utilitarian“ und „incompatible with any recognisably consequentialist ethical theory“ (2005, 244) sei.

dies durch eine weite Pflicht – etwa der Pflicht, den Armen zu helfen – geboten scheint. Konsequentialistische Theorien hingegen betonen typischerweise den Typ von Pflichten, den Kant als ‚weite Pflichten gegen andere‘ bezeichnet.<sup>9</sup>

(ii) *Was ist die Quelle unserer Pflichten?* Auch wenn viele der Handlungen, die Kantianismus und Konsequentialismus vorschreiben, dem Inhalt nach zusammenfallen würden, so sind die Gründe, weshalb sie jeweils vorgeschrieben sind, doch sehr verschieden.<sup>10</sup> Konsequentialisten denken in der Regel, dass wir das Glück oder Wohlergehen aller anstreben sollen, *weil* Glück/Wohlergehen das wichtigste Gut ist. Für Kant sollen wir das Glück aller dagegen nicht primär deshalb befördern, weil es ein basales oder unbedingtes Gut wäre. Vielmehr sei der Grund, warum wir Glück befördern sollen, dass wir andere als Selbstzwecke behandeln müssen, und dazu gehört, dass wir anderen helfen, ihre jeweiligen Zwecke zu erreichen bzw. wir uns deren Zwecke zu eigen machen. Moralische Prinzipien und die Autonomie vernünftiger Akteure sind dem Wert von Glück vorgelegt, bedingen diesen und können ihn einschränken. Konsequentialisten würden Kantianern sicherlich zustimmen, dass es (außer in besonderen Einzelfällen) verboten ist, Akteure für eigennützige Zwecke und zu deren Schaden zu gebrauchen. Der entscheidende Unterschied zwischen Kantianern und Konsequentialisten besteht daher darin, dass Letztere kein strenges Verbot

---

9 In diesem Sinne kritisiert Timmermann (2005, 250), dass Hare alle Pflichten, die nicht den weiten Pflichten gegen andere zuzuordnen sind, vernachlässigt. Siehe auch Dean (2000, 30), der sich kritisch mit Cumminskeys kantischem Konsequentialismus auseinandersetzt.

10 Timmermann (2005) sieht das Hauptargument gegen einen kantischen Konsequentialismus darin, dass aus der Tatsache, dass zwei ethische Theorien in Bezug auf manche konkreten Fälle zu identischen Ergebnissen kommen, nicht geschlossen werden darf, dass sich diese Theorien gleichen. Vgl. auch Hill (2012, 306).

anerkennen, Akteure für die Beförderung des unparteilichen *Guten* zu opfern oder zu schädigen.<sup>11</sup> Es ist ein wichtiger Aspekt des kantischen Instrumentalisierungsverbots, dass sich für den Kantianer das Instrumentalisierungsverbot aus dem besonderen Status des autonomen moralischen Akteurs ergibt bzw. auf diesem gegründet ist. Dies ist wichtig, da es sich als eines der Probleme von Parfits Konzeption herausstellen wird, dass bei ihm der Zusammenhang zwischen Instrumentalisierung, dem Zustimmungsprinzip und Akteurschaft unklar bleibt und sehr lose ist.

## 2.

Nach diesen einführenden Bemerkungen gehe ich nun zu einer kritischen Evaluation des Zustimmungsprinzips über und erläutere zunächst, warum dieses Prinzip in manchen Situationen eine fruchtbare Weiterentwicklung kantischer Gedanken darstellt. Anschließend wende ich mich zwei Kritikpunkten zu: Das Zustimmungsprinzip ist unterbestimmt und daher nicht relevant für die Fälle, die Kantianismus und Konsequentialismus unterscheiden. Zudem lässt das Zustimmungsprinzip, wie Parfit es versteht, mehr Instrumentalisierung zu, als Parfit selbst bereit ist einzuräumen.

Kants Zweck-an-sich-Formel des KI,<sup>12</sup> welche den Status des vernünftigen Akteurs als eines Subjekts, das nicht instrumentalisiert werden darf, zum Ausdruck bringt, hat bekanntermaßen zwei Bestandteile. Sie verlangt zum einen, dass wir an-

---

11 Morgan (2009, 20) sieht einen Hauptvorteil Kants gegenüber dem Konsequentialismus darin, dass für Kant ein vernünftiges Subjekt kein „instrument of impersonal good-making“ sei.

12 „Handle so, daß du die Menschheit, sowohl in deiner Person, als in der Person eines jeden anderen, jederzeit zugleich als Zweck, niemals bloß als Mittel brauchest“ (IV: 429.10–12).

dere niemals nur als Mittel behandeln und, zum anderen, dass wir sie immer auch als Zweck behandeln. Parfit diskutiert die beiden Bestandteile getrennt und ich werde mich auf Letzteren konzentrieren. Die Hauptidee, welche Parfit bezüglich der Behandlung anderer als Zweck diskutiert, ist das *Zustimmungsprinzip*:

It is wrong to treat anyone in any way to which this person could not rationally consent (OWM1, 181).

Wie das Zustimmungsprinzip, laut seiner Theorie, funktioniert, zeigt Parfit am Beispiel *Earthquake*:

*Earthquake*, two people, *White* and *Grey*, are trapped in slowly collapsing wreckage. I am a rescuer, who could prevent this wreckage from either killing White or destroying Grey's leg (OWM1, 185).

Für den Kantianer ist dieser Fall unterbestimmt, da für ihn nur klar ist, dass ich helfen muss – zumindest solange man nicht stipuliert, dass ich versprochen habe, etwas zu tun, was mir keine Zeit zum Helfen lässt, oder sogar versprochen habe, nie zu helfen.<sup>13</sup> In einer Situation, in welcher nur entweder das Retten von Whites Leben oder von Greys Bein möglich ist, steht der Kantianer vor der Schwierigkeit, dass das Prinzip, andere als Zweck an sich zu behandeln, nicht mehr weiter spezifizieren kann, *wem* geholfen werden soll. Der Kantianer kann – zumindest im Rahmen der Zweck-an-sich-Formel<sup>14</sup> – nicht unmit-

---

13 Versprechen zu halten ist, laut Kant, eine strenge Pflicht, helfen eine weite, und erstere Pflichten haben immer Vorrang. Dies kann ganz eigene Probleme für den Kantianismus kreieren, welche ich hier nicht diskutieren kann.

14 Selbst wenn andere Formeln des KI eine klare Lösung vorgäben, stünde der Kantianer damit vor dem Problem, dass er Kants Behauptung,

telbar an die besseren Folgen appellieren, welche das Retten von Whites Leben anstelle von Greys Bein voraussichtlich mit sich bringen würde. Es ist ja gerade einer der entscheidenden Punkte der Zweck-an-sich-Formel, dass sie gewisse Handlungen oder Unterlassungen verlangt aufgrund des Status eines Akteurs als Selbstzweck und unabhängig von den Folgen, welche diese hätten. Steht der Kantianer hier deshalb vor einem moralischen Dilemma, obwohl Kant die Möglichkeit solcher Dilemmata bestreitet?<sup>15</sup> Immerhin können doch sowohl White wie auch Grey meine Hilfe fordern, bzw. beide könnten meiner unterlassenen Hilfe die Zustimmung verweigern. Was das Helfen Whites intuitiv vernünftiger macht, sind scheinbar lediglich die vorhersehbaren Folgen.

Parfits Lösung des Problems basierend auf dem Zustimmungsprinzip ist zumindest für *Earthquake* bestechend. Zwar wäre es für Grey besser und würde seine Wünsche oder Präferenzen besser erfüllen ihn zu retten, das Retten von White ist aber unparteilich besser („make things go in a way that would be impartially better“, OWM1, 186). Parfit weist darauf hin, dass es ihm um *informierte* Zustimmung geht. Er setzt daher voraus, dass die Betroffenen die relevanten Fakten insbesondere in Bezug auf die Folgen der Handlung kennen (OWM1, 184). Laut dem Dualismus der praktischen Gründe hat Grey in *Earthquake* hinreichende parteiliche Gründe, zu verlangen, dass er gerettet wird, weil dies besser für *ihn* ist, wie auch hinreichende unparteiliche Gründe, zu verlangen, dass White gerettet wird, da dies die *unparteilicherweise* bessere Wahl ist. Grey hat sowohl hinreichende Gründe, seine eigene Rettung als auch Gründe,

---

dass die verschiedenen Formeln äquivalent seien (IV: 436.8–9), zurückweisen müsste.

15 Siehe Kant (VI: 224) und Timmermann (2013).

die Rettung Whites zu fordern und diesen zuzustimmen. White hingegen hat hinreichende parteiliche Gründe, zu verlangen, dass ich *ihn* rette, da es für ihn *selbst* besser wäre, gerettet zu werden, sowie hinreichende unparteiliche Gründe, zu verlangen, dass ich ihn rette, da dies *unparteilicherweise* die bessere Wahl ist. Die einzige Option, der *beide* vernünftigerweise zustimmen können oder zu der beide hinreichende Gründe haben, ist, dass Whites Leben gerettet wird (da White keine hinreichenden Gründe hat zuzustimmen, dass Grey gerettet wird).<sup>16</sup>

Parfits Lösung ist kantisch, insofern er das Problem in *Earthquake* nicht einfach als eines von Folgenabschätzungen betrachtet, sondern als eine Frage der praktischen Vernunft. Es geht darum, wozu die von einer Handlung betroffenen Akteure hinreichende Gründe haben zuzustimmen. Dies macht *Earthquake* zu einer Frage der praktischen Vernunft in zwei Hinsichten: Erstens macht es die Frage zu einer, die durch *Gründe* beantwortet werden muss. Zweitens macht es die Frage zu einer, die mit Verweis auf die vernünftigen Vermögen von Akteuren zuzustimmen und Zustimmung zu verweigern und die Achtung vor diesen Vermögen zu beantworten ist. Konsequenzen spielen für diese Frage eine Rolle, da sie ausschlaggebend dafür sind, wie die Dinge unparteilicherweise am besten sind. Parfit zeigt

---

16 Parfit demonstriert dies auch an einem Fall, den er „*Lifeboat*“ nennt: „I am stranded on one rock, and five people are stranded on another. Before the rising tide drowns all of us, you could use a lifeboat to save either me or the five. We are all young, and would lose as much in dying“ (OWM1, 186). Ich habe selbstinteressierte Gründe, dass ich der eine bin, der gerettet wird, aber unparteiliche Gründe, dass die fünf gerettet werden, die fünf haben sowohl selbstinteressierte Gründe als auch unparteiliche Gründe, dass sie gerettet werden. Die einzige Option, zu der wir alle vernünftigerweise zustimmen können, ist, dass die fünf gerettet werden. Parfit betont, dass es in *Earthquake* wie auch *Lifeboat* keinen Unterschied mache, ob die Akteure *de facto* Zustimmung verweigern, solange sie rational zustimmen könnten (OWM1, 192).

hier eine Möglichkeit auf, wie Kantianer Konsequenzen in ihr Denken einbeziehen können: Konsequenzen entscheiden über das unparteiliche Gute und dieses hat einen Einfluss darauf, wozu Akteure vernünftigerweise zustimmen können.

Natürlich wäre hier zu fragen, ob Parfits Interpretation von Kants Zweck-an-sich-Formel exegetisch überzeugend ist. Als Textbeleg dafür, dass das Zustimmungsprinzip eine textlich korrekte Rekonstruktion der Zweck-an-sich-Formel ist, weist Parfit auf IV: 429.33–430.1 hin: „Denn der, den ich durch ein solches Versprechen zu meinen Absichten brauchen will, kann unmöglich in meine Art, gegen ihn zu verfahren, einstimmen und also selbst den Zweck dieser Handlung enthalten.“ Kant selbst spricht hier von Zustimmung als einem wichtigen Kriterium zur Anwendung seiner Formel. Zudem ist Parfits Lesart gegenwärtig sehr verbreitet. Als prominente Vertreter dieser manchmal *possible consent interpretation* genannten Deutung der Zweck-an-sich-Formel nennt Parfit O’Neill (1989), Hill (1992) und Korsgaard (1996). In seiner Diskussion befasst sich Parfit mit gegenwärtigen Kantianern, welche der *possible consent interpretation* anhängen, ähnlich eingehend wie mit Kant. Parfits kritische Diskussion dieses Prinzips ist damit in jedem Fall eine Diskussion einer wichtigen Strömung im gegenwärtigen Kantianismus. Ich möchte im Folgenden offenlassen, ob es andere textlich und systematisch plausible Interpretationen der Zweck-an-sich-Formel gibt, und ich werde im Folgenden die Zweck-an-sich-Formel im Sinne einer *possible consent interpretation* diskutieren. Dies ist natürlich bereits ein großes Zugeständnis an Parfit.

Bevor ich zu einer kritischen Diskussion des Zustimmungsprinzips übergehe, möchte ich noch festhalten, dass zumindest im Rahmen der *possible consent interpretation* Parfit eine wichtige Weiterentwicklung der kantischen Formel bietet.

Kant selbst kann *Earthquake* nicht auf die von Parfit vorgeschlagene Art lösen, da Kant nur reine praktische (moralische) Vernunft, und nichtmoralische (parteiliche) empirische praktische Vernunft kennt. Kant identifiziert jegliche nichtmoralische praktische Vernunft mit Zweckrationalität im Dienste des Eigeninteresses.<sup>17</sup> Er denkt, dass Unparteilichkeit nur in der Form von moralischen Kriterien wie dem kategorischen Imperativ zu haben sei. Parfit führt dagegen eine dritte Option ein: nichtmoralische unparteiliche Vernunft oder die Position der Vernünftigen Unparteilichkeit. Diese zielt auf das, was von einer unparteilichen Perspektive aus die besten Folgen hätte. Parfit hinterfragt Kants angeblich erschöpfende Dichotomie zwischen moralischer Vernunft und parteilicher Selbstliebe und erweitert sie um eine dritte Option. Diese dritte Option ist eine Kategorie, die für Kantianer von genuinem Interesse sein kann. Unparteiliche (nichtegoistische), nichtmoralische Gründe können zumindest in den Fällen, in denen das unparteiliche Gute und unsere moralischen Intuitionen zusammenfallen, helfen, den Zustimmungsgedanken anzuwenden. Diese Kategorie von Gründen verdient, denke ich, eine weitergehende Untersuchung bezüglich ihres möglichen systematischen Stellenwerts für die kantische Philosophie und für den Kantianismus.

---

17 „Alle materiale praktische Principien sind, als solche, insgesamt von einer und derselben Art und gehören unter das allgemeine Princip der Selbstliebe oder eigenen Glückseligkeit“ (V: 22.6–8, siehe auch V: 34.2–11, 35.7–11). Es ist hier nicht klar, ob das allgemeine Prinzip der Selbstliebe und eigene Glückseligkeit zwei verschiedene Prinzipien sind. Kant scheint zu denken, dass sie ein und dasselbe Prinzip unter verschiedenen Namen sind. Dann wäre aber unparteilicher Konsequentialismus, dem es um das Gesamtwohl geht, kein materiales Prinzip, da diese Theorie nicht im Dienste der eigenen Glückseligkeit des Akteurs steht.

Bis hierhin scheint es, dass Parfit Kant fruchtbar weiterentwickelt, indem er zeigt, wie das Zustimmungsprinzip in Fällen wie *Earthquake* intuitiv richtige Resultate ergibt. Wir haben aber bislang noch nicht die entscheidenden Fälle angesprochen, in denen ein Akteur für die Beförderung eines unparteilichen Guten instrumentalisiert wird. Dies sind die Fälle, von denen gemeinhin angenommen wird, dass sie tiefe moralische Meinungsverschiedenheiten zwischen Kantianern und Konsequentialisten konstituieren können.

Parfits Diskussion nimmt eine überraschende Wendung, wenn er auf diese Fälle zu sprechen kommt.

*Means*, White and Grey are trapped in slowly collapsing wreckage. Though White's life is threatened, Grey is in no danger. I could save White's life, but only by using Grey's body as a shield, without Grey's consent, in some way that would destroy her leg. (OWM1, 201)

Man könnte denken, dass Parfit in *Means* wie in *Earthquake* vorschlägt: White hat parteiliche und unparteiliche Gründe zuzustimmen, dass er auf Kosten von Greys Bein gerettet wird, da dies für ihn selbst sowie unparteilicherweise besser wäre, während Grey nur parteiliche Gründe hat, sich nicht zur Rettung Whites instrumentalisieren zu lassen. Damit ist Whites Rettung auf Kosten von Greys Bein die einzige Option, zu der beide vernünftigerweise zustimmen können. Parfit glaubt jedoch, dass es in *Means* moralisch falsch wäre, White auf Kosten von Grey zu retten und er behauptet:

We all have sufficient reasons, I believe, to consent to someone's failing to benefit us, even when this benefit would be as great as the saving of our life, if this way of benefiting us would wrongly injure someone else. (OWM1, 202)

Wir haben nun drei Dinge, auf die wir achten müssen, wenn wir über Zustimmung reden: selbstinteressierte Gründe, unparteiliche Gründe, deontologische Gründe. Letztere ergeben sich daraus, dass etwas moralisch falsch ist.<sup>18</sup>

Ich werde nun zwei Einwände gegen Parfits Zustimmungsprinzip diskutieren. Ich werde argumentieren, dass das Zustimmungsprinzip (i) zu *unterbestimmt* und daher nicht relevant für die Fälle ist, die Kantianismus und Konsequentialismus unterscheiden. Es kann daher nicht dazu dienen zu zeigen, dass keine tiefen moralischen Meinungsverschiedenheiten existieren. Zudem (ii) lässt das Zustimmungsprinzip, wie Parfit es versteht, mehr Instrumentalisierung zu, als er selbst bereit ist einzuräumen. Das Zustimmungsprinzip hat Implikationen, denen er selbst nicht zustimmen würde.

## Problem 1: Unterbestimmtheit

Selbst wenn die Rettung von White moralisch falsch wäre, hieße dies doch nur, dass White und Grey *beide* hinreichende Gründe hätten zuzustimmen, dass White nicht auf Kosten von Grey gerettet wird. Es ist aber weiterhin der Fall, dass Grey und White auch hinreichende Gründe haben, einer Behandlung Greys zuzustimmen, welche White das Leben rettet und Grey instrumentalisiert, da dies *unparteilicherweise* immer noch das Beste wäre. Das Zustimmungsprinzip ist unterbestimmt: Beide Akteure haben sowohl hinreichende Gründe, der Rettung Whites

---

<sup>18</sup> Parfit hat diese Wendung in OWM1, Abschn. 26, bereits angedeutet, wenn er das Zustimmungsprinzip vom Handlungutilitarismus absetzt. Es ist laut Parfit nicht der Fall, dass das Zustimmungsprinzip immer mit dem zusammenfällt, was unparteilich am besten ist: „There are many Utilitarian acts to which some people could not rationally consent“ (OWM1, 191). Problem 2 wird zeigen, dass dies möglicherweise nicht richtig ist bzw. es keinen prinzipiellen Grund gibt, warum dem nach Parfit so sein sollte.

zuzustimmen, als auch, diese abzulehnen. Es scheint, dass das Zustimmungsprinzip die Rettung Whites auf Kosten von Grey gleichermaßen vernünftigerweise verbietet wie fordert.

Situationen, in denen das Zustimmungsprinzip unterbestimmt ist, sind potentiell weit verbreitet. Sie tauchen immer dann auf, wenn es unparteilicherweise besser wäre zu  $\phi$ en, aber  $\phi$ en (intuitiv) moralisch verboten ist (z.B. weil wir Akteure als reines Mittel behandeln oder auf moralisch nicht vertretbare Weise schädigen). Dies sind aber genau die Fälle, die für die Debatte zwischen Kantianismus und Konsequentialismus interessant sind, da Kantianer und Konsequentialisten hier unterschiedliche Urteile abgeben würden. Gemäß der negativen Funktion von OWM2, Schwierigkeiten für eine Vereinigung von Kantianismus und Konsequentialismus auszuräumen, bräuchten wir ein Prinzip für diese Fälle, das beide Positionen, insofern sie vertretbar sind, zufriedenstellt. Dies wäre ein Prinzip, welches uns sagt, in welchen Fällen und warum die eine Position und in welchen die andere recht hat. Parfits Prinzip scheint aber zu implizieren, dass wir hier vor einem Problem stehen, in dem beide Optionen gleichermaßen (un)vernünftig sind. Parfit selbst scheint nicht zu denken, dass das Zustimmungsprinzip im Fall von *Means* unterbestimmt ist, sondern klare Resultate vorgibt.<sup>19</sup> Dies ist aber nicht der Fall, wenn wir Parfits Theorie ernst nehmen. Das Problem der Unterbestimmtheit wird dann besonders deutlich, wenn wir uns vor Augen führen, dass es eine Schwelle der Geringfügigkeit geben muss, welche es moralisch vertretbar macht, Grey zur Rettung Whites auch gegen Greys

19 Parfit erklärt gegen den Einwand, dass das Zustimmungsprinzip in *Means* eine moralisch falsche Handlung verlangt: „But this principle would not, I believe, require this act“ (OWM1, 202). Das Zustimmungsprinzip würde zumindest nicht verlangen, dass wir Grey instrumentalisieren. Es scheint aber, dass das Prinzip dies sowohl verlangt wie auch verbietet.

Willen zu benutzen. Was, wenn Grey sein Bein nicht verliert, sondern nur schwere Schmerzen erleidet? Was, wenn Grey nur leichte Schmerzen oder nur ein leichtes Unwohlsein erleidet? Ein adäquates Prinzip müsste uns helfen zu beurteilen, wann in diesen Fällen keine moralisch verbotene Instrumentalisierung vorläge.

Parfits zentrales Anliegen ist es, die Möglichkeit von tiefen moralischen Meinungsverschiedenheiten auszuräumen bzw. zu zeigen, dass unsere Antworten auf schwierige moralische Fälle nicht davon abhängen, ob wir Kantianer oder Konsequentialisten (oder Kontraktualisten) sind, solange wir eine systematisch vertretbare Form dieser Theorien akzeptieren. Tiefe moralische Meinungsverschiedenheiten scheinen aber durch Parfits Interpretation der Zweck-an-sich-Formel gerade nicht ausgeräumt oder gelöst. Konsequentialisten würden sich in *Means* doch vermutlich auf unparteiliche Gründe für Zustimmung berufen und Kantianer auf deontologische Gründe, Zustimmung zu verweigern. Beide Seiten würden behaupten, die gewichtigere Art von Gründen auf ihrer Seite zu haben. Es ist nicht zu sehen, wie Parfits Zustimmungsprinzip hier einen der zentralen Streitpunkte zwischen Konsequentialisten und Kantianern lösen kann. Anders ausgedrückt: Wenn wir bei Parfits erstem Schritt stehen bleiben, der Einführung unparteilicher Gründe, gewinnen in Zweifelsfällen der Art, wie wir sie diskutiert haben, immer die Konsequentialisten, da wir immer hinreichende Gründe haben, Handlungen zuzustimmen, welche das unparteiliche Gute befördern. Der zweite Schritt, die Einführung deontologischer Gründe, führt hingegen in einen Zustand, in dem sich die entscheidenden Streitfälle nicht vernünftigerweise lösen lassen und in dem tiefe Meinungsverschiedenheiten offenbar werden.

Parfit könnte nun Folgendes erwidern: Das Problem der Unterbestimmtheit, wie gerade beschrieben, ergibt sich nur

dann, wenn das, was den Kantianer dazu veranlassen würde, ein anderes Urteil als der Konsequentialist zu fällen, sich *aus einer Anwendung des Zustimmungsprinzips ergeben muss*. Mit anderen Worten, das Problem der Unterbestimmtheit setzt voraus, dass die Instrumentalisierung von Grey falsch sei, *weil* es gewisse Arten behandelt zu werden gibt, denen vernünftige Wesen nie vernünftigerweise zustimmen können. In diesem Falle wäre es in der Tat so, dass es nicht klar ist, dass Instrumentalisierung in *Means* verboten ist. Zusammen mit unparteilichen Gründen hätten wir dann nämlich die Situation, dass beide Akteure vernünftigerweise beiden Optionen (der instrumentalisierenden wie auch der nichtinstrumentalisierenden) zustimmen können wie auch vernünftigerweise Zustimmung verweigern können.

Möglicherweise meint Parfit dies aber gar nicht. Nach einer alternativen Interpretation *kommt die Falschheit von Instrumentalisierung extern zum Zustimmungsprinzip hinzu, nachdem dieses angewendet wurde*, und übertrumpft dessen Ergebnis.<sup>20</sup> Dann hätte Instrumentalisierung letztlich nichts mit dem Zustimmungsprinzip zu tun, was zunächst einmal gängigen Arten über Instrumentalisierung nachzudenken zuwiderliefe.<sup>21</sup> Zudem würde diese Interpretation Parfit auf eine starke

---

20 Parfit sagt explizit, dass das Zustimmungsprinzip „claims to describe only one of the facts that make acts wrong. So, when this principle does not condemn this way of saving White’s life, it does not thereby *permit* this act, by implying that this act is morally permitted“ (OWM1, 203). Instrumentalisierung könnte also etwas sein, dass Handlungen ganz unabhängig von Zustimmung falsch machen kann.

21 Siehe z.B. O’Neill (1989, 113): „An agent treats another merely as a means and thus wrongly if in his treatment of the other the agent does something to which the other cannot consent“, und Korsgaard (1996, 139): „The question whether another can assent to your way of acting can serve as a criterion for judging whether you are treating her as a mere means“.

Vorrangigkeit moralischer Gründe festlegen, und zwar nicht nur einer Vorrangigkeit von moralischen Gründen vor Eigeninteresse, sondern auch vor *unparteilichen* Gründen. Er würde behaupten, dass, wozu auch immer Akteure vernünftigerweise ihre Zustimmung geben können, es nicht vernünftig sei, Akteure entsprechend zu behandeln, wenn moralische Gründe, die sich aus einem Instrumentalisierungsverbot oder anderen Faktoren ergeben, dagegensprechen. Es ist aber nicht klar, dass, wenn Grey seine vernünftige Zustimmung zur Rettung Whites geben würde, es insgesamt unvernünftig wäre, White auf Kosten Greys zu retten, da moralische Gründe dagegensprechen und diese in jedem Fall überwiegen. Wir würden vermutlich nicht denken, dass es eine moralisch verbotene Instrumentalisierung von Grey darstellt und unvernünftig wäre, White auf Kosten von Grey zu helfen, wenn Grey sich im Vollbesitz seiner geistigen Kräfte dazu entschieden hat, dass White auf Kosten seines eigenen Beins gerettet werden soll.<sup>22</sup> Es scheint inadäquat, in *Means* Instrumentalisierung von Zustimmung zu lösen und als einen Faktor zu be-

---

22 Parfit erkennt explizit an, dass manche Situationen aktuelle Zustimmung erfordern, bzw. vernünftige Zustimmung erfordern kann, dass aktuelle Zustimmung vorliegt (OWM1, Abschn. 27). In *Earthquake* wie auch in *Lifeboat* mache es aber keinen Unterschied, ob die Akteure *de facto* Zustimmung verweigern, solange sie rational zustimmen können (OWM1, 192). Interessanterweise diskutiert Parfit die genaue Rolle von aktueller Zustimmung nicht in Zusammenhang mit *Bridge* und *Transplant*. Er argumentiert hier sogar, dass das Zustimmungsprinzip eine Handlung nicht verbietet, wenn der Akteur vernünftigerweise zustimmen kann, aber aktuelle Zustimmung fehlt. Solche Handlungen, da sie vom Zustimmungsprinzip erlaubt sind, stellen keine verbotenen Instrumentalisierungen dar. Parfit schließt: „It might be wrong for you to kill me, without my consent, as a means of saving the five. But that is not implied by these Kantian principles“ (OWM1, 221). Hier haben wir wieder den Gedanken, dass wenn Instrumentalisierung falsch ist, dies nichts mit dem Zustimmungsprinzip zu tun hat.

handeln, der extern hinzukommt. Was in *Means* den Ausschlag gibt, ist Zustimmung bzw. deren Ausbleiben.

Die Instrumentalisierung Greys in *Means* ist nur aufgrund einer zusätzlichen Stipulation von Seiten Parfits verboten. Das Manöver, das Zustimmungsprinzip an unsere Intuitionen anzulehnen, scheint eine *Ad-hoc*-Stipulation, welche das Prinzip notdürftig reparieren soll. Parfit führt zusätzliche Gründe ein, die sich daraus ergeben, dass etwas intuitiv falsch sei. Dies ist fragwürdig, da wir doch erwarten, dass die Zweck-an-sich-Formel *selbst* bzw. das Zustimmungsprinzip von der Falschheit der Instrumentalisierung oder von dem, was das Retten von White auf Kosten Greys in *Means* falsch macht, ausgeht bzw. eine philosophische Systematisierung dieser Falschheit darstellt. Hier wirkt es aber so, als käme die Falschheit von Instrumentalisierung noch extra zu dem Prinzip hinzu.

Entweder ist das Zustimmungsprinzip in Fällen wie *Means* unterbestimmt und ergibt nicht das Resultat, von dem Parfit meint, dass es intuitiv richtig ist, oder es hat nichts mit auf Instrumentalisierung beruhender moralischer Falschheit zu tun und diese muss von Parfit zusätzlich eingeführt werden. Wenn Parfit nichts anbieten kann, um Streitfälle wie *Means* auf Basis kantischer wie auch konsequentialistischer Einsichten, insofern diese Einsichten jeweils vertretbar sind, zu lösen, müssen wir davon ausgehen, dass in diesen Fällen zwischen Kantianismus und Konsequentialismus tiefe moralische Meinungsverschiedenheiten bestehen.

## Problem 2: Organraub

Dass es Parfit mit seinem Zustimmungsprinzip letztlich nur um unparteiliche Gründe geht, nicht um das, was Kant in seiner Zweck-an-sich-Formel thematisiert, nämlich den besonderen

moralischen Status autonomer Akteure, wird deutlich an einem weiteren Beispielfall, den Parfit diskutiert:

*Bridge*: The train is headed for the five, but there is no other track and tunnel. I am on a bridge above the track. Your only way to save the five would be to open, by remote control, the trap-door on which I am standing, so that I would fall in front of the train, thereby triggering its automatic brake (OWM<sub>1</sub>, 218).

Parfit sagt dazu:

If the choice were mine, I would have sufficient reasons to jump in front of the train, so that it would kill me rather than the five. [...] Since I could rationally kill myself as a means of saving the five, I could also rationally consent to your treating me in this way.

It might be objected that I could not rationally consent to your killing me as a means, because this act would be wrong. But if I consented to this act, it would not be wrong. So even if this act would be wrong without my consent, that would not give me any reason to refuse consent (OWM<sub>1</sub>, 220).

In *Means* war es letztlich eine unbefriedigende *Ad-hoc*-Maßnahme Parfits, die das Zustimmungsprinzip mit kantianischen Überzeugungen in Einklang brachte. Ich habe bereits argumentiert, dass es in *Means* einen Unterschied machen muss, ob Grey der Rettung Whites zustimmen kann. Parfit scheint dies auch so zu sehen. In *Bridge* können wir daher Parfit, laut seinen eigenen Aussagen, vor den Zug werfen, da er hinreichende unparteiliche Gründe hat, vernünftigerweise dieser Behandlung zuzustimmen, und laut Zustimmungsprinzip nichts gegen diese Behandlung Parfits spricht, da Akteure immer hinreichende

Gründe zu haben scheinen, einer Behandlung zuzustimmen, welche das unparteilich Gute so weit wie möglich befördert.

Wenn tatsächliche Zustimmung zu einer Behandlungsweise vorliegt, dann stellt diese Behandlungsweise nach Parfit entweder keine Instrumentalisierung dar oder, wenn doch, dann ist diese keine Instrumentalisierung in einer Form, die hinreichende Gründe gegen diese Behandlungsweise liefern kann. Nach dem Zustimmungsprinzip kommen konsequentialistische Überlegungen über die Theorie nichtmoralischer Gründe in moralische Überlegungen herein, und angewendet auf Einzelfälle ergeben sich die Resultate, welche konsequentialistische Theorien verlangen würden. Das vermeintlich kantische Zustimmungsprinzip scheint eher ein Instrumentarium zu sein, das – orientiert an den besten Folgen – der Anwendung unparteilicher Gründe auf Einzelfälle dient. Dies wird am deutlichsten an dem, was laut Zustimmungsprinzip bei einem weiteren Beispielfall, *Transplant*, herauskommt, gegenüber dem, was laut Parfit dabei herauskommen *sollte*:

*Transplant*, I am in hospital, to have some minor operation. You are my doctor. You know that, if you secretly killed me, my transplanted organs would be used to save the lives of five other people (OWM1, 363).

In dieser Situation besteht aus konsequentialistischer Sicht eine große Versuchung, dem Arzt zu sagen, dass er mich ausweiden muss,<sup>23</sup> wohingegen der Kantianer dies als streng verboten betrachten würde. Würden die Urteile von Kantianismus und Konsequentialismus derartig auseinanderfallen und wären sie in vollkommen unterschiedlichen Vorstellungen von dem, was in solchen Situationen den Ausschlag gibt, gegründet, dann

---

23 Ein *locus classicus* hierfür ist Harris (1975).

bestünde wenig Hoffnung, Kantianismus und Konsequentialismus in einer einheitlichen Theorie zu vereinigen.

Im Falle von *Transplant* folgt Parfit einem bekannten regelkonsequentialistischen Muster und argumentiert mittels versteckter Kosten. Parfit argumentiert, dass, wenn wir wüssten, dass immer wenn wir in einem Krankenhaus sind, ein (geringes) Risiko besteht, dass unsere Organe entnommen werden, dies das Vertrauen in Ärzte und das Gesundheitssystem so schwächen würde, dass eine allgemeine Regel, Patienten auszuweiden, die Welt nicht besser machen würde. Parfit nennt dies das „*Anxiety and Mistrust Argument*“ (OWM1, 363). Gegen Parfit ist einzuwenden, dass es eine empirische Frage ist und keineswegs ausgemacht, ob gelegentlicher Organraub in Krankenhäusern das Vertrauen in das Gesundheitssystem so weitgehend unterminieren würde, dass die Folgen schlecht sind. Organraub könnte immerhin auch dazu führen, dass das Vertrauen in das Gesundheitssystem steigt, da Akteure wüssten, dass, wenn sie aufgrund schwerer Erkrankungen oder Unfälle neue Organe brauchen, das Gesundheitssystem alles tun wird, ihnen zu helfen. Zudem ist es statistisch sehr viel wahrscheinlicher, dass ich vom Organraub profitiere, als dass ich durch ihn Schaden erleide.

Weiterhin setzt Parfit voraus, dass der Organraub *nicht heimlich* geschehen kann, sondern als Regel der Öffentlichkeit bekannt sein muss bzw. dass wir die Regel anhand der Frage evaluieren müssen, welche Prinzipien niemand vernünftigerweise zurückweisen kann. Dafür müssen wir berücksichtigen, welche Auswirkungen es hätte, wenn jeder diese Prinzipien akzeptiert (OWM1, 364). Dies scheint aber eine unnötige Zusatzannahme von Seiten Parfits zu sein. Ein Philosoph kann sich fragen, ob Akteure vernünftigerweise ein Prinzip akzeptieren oder zurückweisen *würden*, von dem diese Akteure *de facto*

nichts wissen.<sup>24</sup> Würden im Fall von *Transplant* fünf getötet und nur eine Person gerettet, wäre es, *ceteris paribus*, klar, dass jeder gute Gründe hat, das Prinzip zurückzuweisen, selbst wenn er nichts von ihm wüsste. Es ist nicht klar, warum Parfit hier darauf beharrt, dass das Prinzip des Organraubs öffentlich gemacht werden muss, oder warum der Organraub nicht eine Spezifizierung einer allgemeineren und scheinbar harmlosen öffentlichen Regel sein kann, wie etwa: Das Gesundheitssystem soll so viele Menschen wie möglich retten. Letztere Regel würde vermutlich allgemeine Zustimmung erhalten, aber möglicherweise auch gelegentlichen Organraub erlauben. Die meisten AkteurInnen wüssten aber nicht, dass Organraub eine Folge der öffentlichen Regel ist, die sie befürworten.<sup>25</sup> Parfits Zurückweisung des Organraubs kann nicht überzeugen. Das Argument über versteckte Kosten ignoriert den zentralen Umstand, dass es sich beim Organraub um eine moralisch verbotene Instrumentalisierung handelt. Das Instrumentalisierungsverbot bzw. die Forderung, Akteure zu respektieren und nicht gegen deren Willen zu benutzen, scheint eine intuitiv einleuchtendere Begründung des Organraubverbots darzustellen als die Überlegungen, die Parfit selbst anführt.<sup>26</sup>

---

24 Parfit benutzt diese Idee selbst: Wir können einen Akteur täuschen, wenn es schlimme Folgen für den Akteur hätte, wenn er die Wahrheit wüsste. Dies ist legitim, da der Akteur, wenn wir ihm die Wahrheit erzählen *würden*, zugestimmt *hätte*, dass wir ihn täuschen können (OWM1, 179).

25 Der Hintergrund für Parfits Zusatzannahme, dass die Regel öffentlich sein muss, ist vermutlich, dass Parfit hier Scanlons Kontraktualismus diskutiert. In seiner Diskussion des Zustimmungsprinzips betont Parfit, dass es ihm weniger um die Zustimmung zu allgemeinen Regeln geht, sondern um Zustimmung zu bestimmten Handlungen (OWM1, 182).

26 Selbst wenn meine Argumente gegen Parfits Behandlung von *Transplant* nicht überzeugen, kann man immer noch einen Fall stipulieren,

Wir können nun für *Transplant* analog zu *Bridge* die oben zitierte längere Parfit-Passage (OWM1, 220) wie folgt abwandeln:

Wenn die Wahl meine wäre, hätte ich hinreichende (unparteiliche) Gründe, mich töten zu lassen und meine Organe unter den fünf aufteilen zu lassen. Da ich vernünftigerweise zustimmen kann, als Mittel zur Rettung der fünf getötet zu werden, kann ich auch vernünftigerweise zustimmen, von dir auf diese Art behandelt zu werden. Es könnte eingewendet werden, dass ich nicht vernünftigerweise zustimmen kann, als Mittel von dir getötet zu werden, da diese Handlung (das Töten als Mittel) falsch wäre. Aber wenn ich dieser Handlung dennoch zustimmte, wäre sie demnach nicht falsch. Sogar wenn diese Handlung ohne meine Zustimmung falsch wäre, würde mir dies keine Gründe geben, Zustimmung zu verweigern.

Es scheint, dass in *Transplant* wie auch in *Bridge* unparteiliche Gründe es für mich vernünftig machen zuzustimmen, geopfert zu werden, und dass das Zustimmungsprinzip dieses Opfer

---

in welchem die versteckten Kosten nicht vorliegen: Ein Medikament, das voraussichtlich viele Leben retten wird, muss vor Einführung an einigen gesunden Menschen getestet werden, um die chemische Zusammensetzung für menschliche Bedürfnisse zu optimieren. Es ist zwar äußerst wahrscheinlich, dass diese Optimierung gelingen wird, jedoch werden ebenso mit hoher Wahrscheinlichkeit einige der Testpersonen schlimme Folgeschäden erleiden. Es ist daher nicht möglich, freiwillige Testpersonen zu finden. Die medizinische Entwicklung der Menschheit ist bereits so weit fortgeschritten (es gibt Heilmittel gegen fast alle Krankheiten), dass es voraussichtlich keine ähnlichen Situationen mehr geben wird (siehe Hoesch, Sticker [im Erscheinen] für diesen Fall). Der Kantianer würde sagen, dass die Instrumentalisierung von Individuen hier verboten ist, und Parfit kann nicht einfach sagen, dass er aufgrund von versteckten Kosten zum gleichen Ergebnis kommt.

nicht verbieten kann. Es kann allenfalls sagen, dass ich Gründe habe, mich opfern zu lassen, wie auch mich nicht opfern zu lassen. Parfits Zustimmungsprinzip und der Umstand, dass die Schlechtheit von Instrumentalisierung nicht notwendig mit diesem Prinzip verbunden ist, legen in *Transplant* ein Urteil nahe, das letztlich sogar mit Parfits eigenen Vorstellungen konfligiert bzw. zu konsequentialistisch für Parfit ist. Woran Kantianer dagegen festhalten wollen, ist, dass das Opfern eines vernünftigen Akteurs verboten ist *aufgrund des besonderen Status autonomer Vernunftwesen*, und zwar für jeglichen Zweck, auch zur Beförderung *unparteilicher* Zwecke. Wenn Parfit so etwas nicht anbieten kann, ist es fraglich, inwieweit zentrale kantische Elemente bei ihm überhaupt vorkommen, ob er diese in seine Theorie aufgenommen hat und ob er tiefe moralische Meinungsverschiedenheiten zwischen Kantianern und Konsequentialisten ausräumen kann. Parfit scheint eine konsequentialistische Position vorauszusetzen und er kann daher die zentralen kantischen Elemente, die er berücksichtigen müsste und auch berücksichtigen möchte, nicht aufnehmen.

Parfit schließt an seine Diskussion des Zustimmungsprinzips eine Reihe von Bemerkungen dazu an, was es heißt, Akteure als reines Mittel zu behandeln (OWM, Kap. 9). Ich kann auf diese hier nicht näher eingehen, zumal Parfit einräumt, dass die Anweisung, Akteure nicht als reine Mittel zu behandeln, nur etwas darüber sagt, wie wir Akteure *betrachten* sollen. Auf der Handlungsebene gibt uns dieses Prinzip keine Anweisungen, die wir nicht auch durch andere Prinzipien (wie etwa das Nichtschädigungsgebot) erhalten (OWM1, 232).

Wir können abschließend festhalten, dass Parfits Vorschlag, unparteiliche Gründe in die Debatte um Zustimmung einzubringen, in den entscheidenden Fällen (unparteiliches Gu-

tes vs. moralische Falschheit) unterbestimmt ist. Es kann daher hier nicht davon gesprochen werden, dass Kantianismus und Konsequentialismus bzw. die Intuitionen oder Überzeugungen, auf welche sich diese Theorien stützen, wirklich vereinbar gemacht werden. Zudem scheint die moralische Falschheit von Instrumentalisierung losgelöst zu sein von dem eigentlichen Prinzip, welches Parfit vorschlägt, und damit begründungstechnisch im luftleeren Raum zu hängen.

## Konklusion

Parfits Diskussion der Zweck-an-sich-Formel und speziell des Zustimmungsprinzips zeigt Schwierigkeiten auf, Kantianismus und Konsequentialismus einander anzunähern. Diese Schwierigkeiten bleiben weitgehend ungelöst. In Teil 2 seines Buches gelingt es Parfit weder in Einzelfällen noch auf einer theoretischen Ebene, die Zweck-an-sich-Formel überzeugend mit konsequentialistischen Gedanken in Einklang zu bringen. Es zeigte sich, dass es tiefe moralische Meinungsverschiedenheiten zwischen der kantischen Zweck-an-sich-Formel und konsequentialistischen Theorien geben kann, wenn diese z.B. verlangen würden, eine Person in *Bridge* vor die Schienen zu werfen oder ihr in *Transplant* Organe zu entnehmen und zu verteilen, aber die Zweck-an-sich-Formel dies verbietet. Parfit will keinen Organraub, aber seine Prinzipien sowie seine Diskussion dieser Prinzipien zeigen, dass seine Theorie ein prinzipielles Verbot des Organraubs nicht hergeben. Der Verweis auf versteckte Kosten überzeugt nicht, Parfits Diskussion von *Bridge* impliziert, dass Organraub erlaubt ist und das Zustimmungsprinzip scheint dies zumindest nicht verbieten zu können. Parfits Zustimmungsprinzip betrachtet jene Fälle, in denen Kantianer und Konsequentialisten verschieden urteilen würden, entweder als

unauflöslich bzw. nur durch Zusatzannahmen, die weitgehend *ad hoc* sind, auflösbar (Problem 1) oder als immer zu Gunsten des unparteilich Guten auflösbar – sogar dann, wenn Parfit selber denkt, dass dies unangemessen ist (Problem 2). Es ist nicht zu sehen, wie diese an internen Spannungen reiche Konzeption dazu beitragen kann, Dissens zwischen kantischen und konsequentialistischen Ideen auszuräumen.

### Literatur

- Cummiskey, David (1990): „Kantian Consequentialism“, *Ethics*, vol. 100 (3), 586–615.
- (1996): *Kantian Consequentialism*, Oxford University Press.
- Dean, Richard (2000): „Cummiskey’s Kantian Consequentialism“, *Utilitas* 12 (1), 25–40.
- Hare, R. M. (1993): „Could Kant Have been A Utilitarian?“ *Utilitas* 5 (1), 1–16.
- Harris, John (1975): „The Survival Lottery“ *Philosophy* 50, 81–87.
- Hill, Thomas E. (1992): *Dignity and Practical Reason*, Cornell University Press.
- (2012): *Virtue, Rules, and Justice. Kantian Aspirations*, Oxford University Press.
- Hoesch, Matthias, Sticker, Martin (im Erscheinen): „Parfit über Kantianismus und Konsequentialismus“, in: Hoesch, Munders, Rütter (Hgs.) *Derek Parfit in der Diskussion*, Meiner: Hamburg.
- Kant, Immanuel (1900 ff.): *Kant’s gesammelte Schriften*, herausgegeben von der Königlich Preußischen Akademie der Wissenschaften, Georg Reimer: Berlin.
- Korsgaard, Christine (1996): *Creating the Kingdom of Ends*, Cambridge University Press.
- Morgan, Seiriol (2009): „Can there be a Kantian Consequentialism?“, *Ratio* 22 (1), 19–40.
- O’Neill, Onora (1989): *Constructions of Reason*, Cambridge University Press.
- Parfit, Derek (2011): *On What Matters*, Oxford University Press.

- 
- Timmermann, Jens (2005): „Why Kant Could not Have Been a Utilitarian“, *Utilitas*, vol. 17 (3), 243–264.
- (2013): „Kantian Dilemmas? Moral Conflict in Kant’s Ethical Theory“, *Archiv für Geschichte der Philosophie*, vol. 95 (1), 36–64.
- Wood, Allen (1998): „Humanity as End in Itself“, in: Guyer (ed.): *Kant’s Groundwork of the Metaphysics of Morals. Critical Essays*, Rowman & Littlefield: Lanham et. al., 165–189.

