

# Personen, Normativität, Moral: Eine Einführung in die Philosophie Derek Parfits

Derek Parfit (1942-2017) gehörte nicht zu den Intellektuellen, die in der Öffentlichkeit viel von sich reden machen. Vielmehr erinnert sein Wirken in mancher Hinsicht fast schon an das Urbild des Wissenschaftlers im Elfenbeinturm: Seine jahrzehntelange Position als Fellow in Oxfords All Souls College hat es ihm gestattet, seine Kräfte unermüdlich auf die philosophische Forschung und den Austausch mit den Fachkollegen zu konzentrieren. Ganz am Erkenntnisgewinn orientiert, war Parfit an Medienpräsenz wenig interessiert.<sup>1</sup> In den akademischen Debatten zur praktischen Philosophie der Gegenwart ist er dafür allgegenwärtig. Seine Hauptwerke *Reasons and Persons* (im Folgenden: RaP) und *On What Matters* (im Folgenden: OWM) gehören sowohl in der internationalen Debatte als auch im deutschen Universitätsalltag ebenso zum Kanon der wichtigen Primärwerke wie zahlreiche seiner Aufsätze. Seine Art, sich ohne Umwege und zugleich bis in die tiefsten Verzweigungen vieler zentraler Probleme der Philosophie anzunehmen, macht ihn über seinen Tod hinaus zu einem unverzichtbaren Diskussionspartner.

Parfits Leben lässt sich in wenigen Sätzen zusammenfassen: 1942 in China als Sohn eines britischen Ärzte-Ehepaars geboren, das sich der Arbeit in Missionskrankenhäusern widmete, studierte Parfit in Oxford zunächst Geschichte. 1965/66 entdeckte er bei einem Aufenthalt an der Columbia University und in Harvard die Philosophie für sich und wurde bereits 1966 Fellow am All Souls College in Oxford, wo er – von einigen Gastprofessuren abgesehen – den Rest seiner philosophischen Karriere verbrachte, bis er am 1. Januar 2017 überraschend verstarb.

Der philosophische Durchbruch gelang Parfit auf breiter Linie spätestens 1984 mit dem bereits nach wenigen Jahren zum Klassiker

1 Aus den wenigen an eine breitere Öffentlichkeit adressierten Beiträgen ragt eine 2011 im *New Yorker* erschienene Reportage hervor, die zugleich die umfangreichste biographische Darstellung zu Parfit enthält (MacFarquhar 2011). Auf Deutsch erschien im Philosophiemagazin *Hohe Luft* 2013 ein ausführliches Interview (Parfit 2013), in der FAZ 2017 ein Nachruf (Dath 2017).

avancierten Buch *Reasons and Persons*.<sup>2</sup> Neben Überlegungen zur Theorie rationaler Entscheidung sowie zu Problemen des Konsequentialismus und der *common sense morality* enthält dieses Buch vor allem eine Theorie personaler Identität und Thesen zu unseren Pflichten gegenüber zukünftigen Generationen.

Im Anschluss widmete Parfit sich lange der Suche nach der plausibelsten Moralthorie und ihren metaethischen Implikationen. Seine Ergebnisse stellt er zunächst 2002 im Rahmen der *Berkeley Tanner Lectures on Human Values* vor. Nachdem das Manuskript jahrelang unter vielen Kollegen im englischsprachigen Raum zirkuliert hatte und durch Parfits Reaktionen auf die vorgebrachte Kritik immer umfangreicher geworden war, erschienen 2011 zwei Bände unter dem Titel *On What Matters*. Während Band 1 von einer Auseinandersetzung mit dem Kantianismus dominiert ist und in der Formulierung der sogenannten *Triple Theory* gipfelt, enthält Band 2 einerseits kritische Kommentare von Susan Wolf, Allen Wood, Barbara Herman und Thomas Scanlon, die überwiegend im Rahmen der *Tanner Lectures* vorgetragen wurden;<sup>3</sup> und andererseits Parfits eigene umfangreiche Überlegungen zur Metaethik. Mittlerweile ist, angeregt durch kritische Einwände in dem von Peter Singer herausgegebenen Buch *Does Anything Really Matter?*, ein dritter Band erschienen, der insbesondere Parfits Thesen zur Metaethik weiterführt.

Die von Parfit behandelten Themen reichen vom Personenbegriff über Fragen der Normativität im Allgemeinen – also der Klärung des Status von normativen Gründen überhaupt, solche des Eigeninteresses eingeschlossen – bis hin zu grundlegenden Fragen der Moral. »Person«, »Normativität« und »Moral« sind daher die Schlagwörter, mit denen sich das Themenspektrum Parfits umschreiben lässt.

Dabei sind seine Argumente nicht nur schon für sich genommen komplex, sondern hängen auch in ganz unterschiedlicher Form zusammen: Während die vertretenen Thesen teils aufeinander aufbauen, werden zuweilen Überlegungen bewusst unabhängig von den andernorts erzielten Ergebnissen eingeführt und begründet oder gar gezielt aus konkurrierenden Prämissen abgeleitet.<sup>4</sup>

2 Schon zuvor hatte Parfits Aufsatz »Personal Identity« von 1971 bereits einiges Aufsehen erregt.

3 Nur Barbara Hermans Beitrag geht nicht auf die *Tanner Lectures* zurück.

4 So baut etwa die normative Ethik nicht auf der Theorie personaler Identität

Auf diese Weise gelingt es Parfit, mit Teilaspekten seiner Theorie auch Philosophen zu überzeugen, die manche seiner fundamentalen Überzeugungen ablehnen. Die nachstehende Darstellung von Parfits Denken muss deshalb nicht nur zahlreiche Aspekte ausklammern, sondern kann auch keine in sich geschlossene Theorie präsentieren. Sie nimmt ihren Ausgang bei der Theorie personaler Identität und ihren ethischen Implikationen (Abschnitt 1), behandelt daraufhin – grundlegende Annahmen der Gründe-Theorie einschließlich – die normative Ethik Parfits (Abschnitt 2) und stellt schließlich das metaethische Fundament seiner Moralphilosophie dar (Abschnitt 3).

## 1. Die Theorie der personalen Identität

Als Parfit in den 1970er Jahren begann, sich mit dem Thema der personalen Identität auseinanderzusetzen, war der Personenbegriff bereits ein konsolidiertes und kanonisiertes Thema auf der philosophischen Agenda. Die Gründe für diese zentrale Stellung des Personenbegriffs innerhalb der analytischen Philosophie sind vielfältig. Ein Grund ist sicher, dass er mit zentralen Fragen unseres menschlichen Selbstverständnisses verbunden ist.<sup>5</sup> Es ging oft nicht um den Begriff der Person selbst, sondern man war optimistisch, dass eine Analyse dessen, was eine Person ist, uns zum Beispiel Aufschluss darüber gibt, wie das Leib-Seele-Verhältnis zu verstehen ist, wie sich das Rätsel der Willensfreiheit lösen lässt oder wie die Mensch-Tier-Unterscheidung konzipiert werden muss.

Den Diskussionskontext für Parfit bilden dabei vor allem zwei Arbeiten, die in ihrer Ausrichtung ganz unterschiedlich sind – Pe-

auf, auch wenn beide sich in mancher Hinsicht zu einem passenden Gesamtbild ergänzen; und die *Triple Theory* ist weder auf den Prioritarismus noch auf eine Lösung der sogenannten »Abstoßenden Schlussfolgerung« festgelegt. Selbst innerhalb von OWM gibt es zahlreiche Exkurse mit Argumentationsstrategien, anhand deren die Kernthesen für den Fall verteidigt werden können, dass jemand andere Prämissen zugrunde legt als Parfit selbst.

5 Oder wie es Harry Frankfurt ausdrückt: »[D]ie Kriterien für das Personsein sind [...] so gestaltet, dass sie diejenigen Merkmale erfassen, die Gegenstand unserer tiefgreifendsten menschlichen Sorge um uns selbst sind und die Quelle desjenigen, was wir als das Bedeutsamste und Herausforderndste in unserem Leben betrachten.« (Frankfurt 1971, S. 12; Übers.: Hg.)

ter Strawsons *Individuals* (1959) und Harry Frankfurts *Freedom of the Will and the Concept of a Person* (1971). Auf der einen Seite Strawson, der die zu seiner Zeit vorherrschenden naturalistischen und cartesischen Paradigmen kritisiert, die ihm zufolge Personen entweder als primär körperliche Wesen oder als geistige Entitäten verstehen. Personen seien aber weder als »beseelte Körper« noch als »körperliche Seelen« zu verstehen, sondern als schlicht nicht weiter analysierbare primitive Entitäten.<sup>6</sup> Auf der anderen Seite Frankfurt, der die Unanalysierbarkeitstheese von Strawson scharf kritisiert: Diese führe nach ihm dazu, dass wir einen uninformativen, weil nur negativ bestimmbareren Personenbegriff erhielten. Positiv übersehe Strawson die Möglichkeit, dass wir Personen durchaus analysieren können – nämlich als Wesen, die durch ihr spezifisches Wollen und die Möglichkeit der Bezugnahme auf sich selbst definiert sind.

In diese Debatte hinein unterbreitet Parfit seinen eigenen Vorschlag, den er samt seiner Implikationen in mehreren Anläufen darstellt. Eine erste Fassung findet sich bereits 1971 in seinem Aufsatz »Personal Identity«,<sup>7</sup> eine ausführliche Ausarbeitung erfährt sie aber erst im dritten Teil von RaP, insbesondere im zwölften Kapitel, »Why Our Identity Is Not What Matters«. Seitdem haben sich weitere Modifikationen ergeben. Die einschlägigste Zusammenfassung stellt der in diesem Band abgedruckte Beitrag »The Unimportance of Identity« von 1995 dar.

Methodisch orientiert sich Parfit an der Art von Begriffsexplikation, die seit den 1970er Jahren Einzug in die analytische Philosophie gehalten hat. Markant ist allerdings sein Vorgehen, unsere verbreiteten Vorannahmen über den Personenbegriff zu testen und auf ihre Haltbarkeit zu prüfen. Parfit bezieht sich hierbei auf fiktive Beispiele, welche es ihm erlauben, einzelne Aspekte unseres Personenverständnisses zu betonen, ohne auf die Komplexität der Gesamtsituation eingehen zu müssen – ein Vorgehen, das ihm nicht nur Lob eingebracht hat.<sup>8</sup>

6 Für beide Ausdrücke vgl. Strawson 1959, S. 103.

7 Eine deutsche Übersetzung dieses älteren Textes findet sich in Quante 1999, S. 71-99.

8 Vgl. dazu etwa die Kritik von Kathleen Wilkes, für die die Methode der Gedankenexperimente eine grundsätzlich falsche Herangehensweise ist, um eine Theorie der Person zu entwickeln. Vgl. etwa ihr Buch *Real People* (1988), welches bezeichnenderweise den Untertitel *Personal Identity without Thought Experiments* trägt.

Inhaltlich wendet sich Parfit dabei nicht – wie noch Strawson und Frankfurt vor ihm – in synchroner Perspektive direkt dem Personenbegriff zu, sondern fragt zunächst nach den Bedingungen der diachronen (oder gleichbedeutend: transtemporalen) Identität: Welche Bedingungen müssen erfüllt sein, damit die durch den singulären Term  $A$  zum Zeitpunkt  $t_1$  bezeichnete Person identisch ist mit der durch den singulären Term  $B$  zum Zeitpunkt  $t_2$  bezeichneten Person?

### 1.1 Die These der Bedeutungslosigkeit der Identität

Trotz aller Modifikationen im Detail sind Parfits wesentliche Gedanken über die Zeit hinweg dieselben geblieben.<sup>9</sup> Sie lassen sich anhand der im akademischen Diskurs bereits zum Slogan geronnenen Bedeutungslosigkeitsthese zusammenfassen: »Identity is not what matters in survival.«

Um Missverständnisse auszuschließen, sind zwei Hinweise unerlässlich:

Zum einen ist es wichtig, sich von der indikativischen Formulierung nicht täuschen zu lassen. Parfit möchte nicht nur ein ontologisch-metaphysisches Problem lösen, welches die Einheit der Person betrifft. Vielmehr geht es ihm auch darum, eine bestimmte Haltung zu diesem Problem zu empfehlen. Parfit behauptet nicht lediglich, dass Identität nicht dasjenige *ist*, worauf es beim Weiterleben ankommt. Vielmehr verbindet er den Slogan mit einer normativen Aufforderung zur Verhaltensänderung, die gegen den sich am Common Sense orientierenden philosophischen Mainstream gerichtet ist: Personale Identität *sollte* – anders als es generell getan wird – nicht dasjenige sein, worauf es uns ankommt, wenn wir weiterleben.

Zum anderen besagt die Bedeutungslosigkeitsthese weder, dass die Rede von personaler Identität sinnlos sei, noch, dass eine solche Identität niemals vorliegen könne. Tatsächlich wendet sich Parfit gegen eine bestimmte Art, die Frage nach der personalen Identität anzugehen: Eine bestimmte Identitätskonzeption sei bei der Suche nach einheitsstiftenden Kriterien für diachrone Identität rundum abzulehnen. Hierbei bilden die *Simple Views* seinen Hauptangriffs-

<sup>9</sup> Eine Ausnahme stellt die Neueinführung eines physischen Kriteriums für diachrone Identität dar; siehe Abschnitt 1.3.

punkt. Diese gehen – nach Parfit fälschlicherweise – davon aus, dass die diachrone Identität in der nicht weiter analysierbaren Identität eines Selbst besteht. Worin aber besteht ein solches »Selbst«? Parfit zufolge vertreten die *Simple Views* eine ontologische Interpretation, die das »Ich« als (materielle oder immaterielle) Substanz im Sinne eines eigenständigen Trägers von Eigenschaften deutet. Wesentlich für die diachrone Identität einer Person ist dann die Erhaltung dieses Trägers, etwa des Körpers oder der Seele, über die Zeit hinweg. Das ist eine sehr wirkmächtige Position, die viele Anhänger gefunden hat – von den traditionellen Materialisten über die Cartesianer bis zu den meisten Philosophen zur Zeit der Entstehung von Parfits Konzeption. Es ist daher nicht verwunderlich, dass Parfits Vorschlag als Provokation aufgefasst wurde.

## 1.2 Die lockeanische Theorie der diachronen Identität

Was sind nun die Gründe für Parfits Ablehnung der *Simple Views*? Für ihn liegt der Hauptgrund darin, dass er einen eigenständigen substanzontologischen Träger nicht für notwendig hält, um die transtemporale Einheit der Person sicherzustellen. Um diese negative These zu belegen, werden die verschiedensten Gedankenexperimente vorgetragen. Eines der prominentesten findet sich zu Beginn von »The Unimportance of Identity«:

Beginnen wir mit einem Science Fiction-Szenario. Hier auf der Erde betrete ich einen Teletransporter. Wenn ich einen Knopf drücke, zerstört eine Maschine meinen Körper und registriert dabei den genauen Zustand all meiner Zellen. Diese Informationen werden per Funk zum Mars gesendet, wo eine andere Maschine aus organischem Material eine perfekte Kopie meines Körpers herstellt. Die Person, die auf dem Mars aufwacht, scheint sich daran zu erinnern, mein Leben bis zu dem Zeitpunkt, an dem ich den Knopf drückte, gelebt zu haben, und ist auch in jeder anderen Hinsicht genau wie ich.<sup>10</sup>

Dieses Szenario soll nach Parfit zeigen, dass es logisch möglich ist, die Existenz der Kopie auf dem Mars als Weiterleben der fraglichen

10 In diesem Band, S. 69.

Person zu werten – und zwar, ohne dass ein »Ich« oder »Selbst« als Träger von Eigenschaften eine Rolle bei der Erklärung spielt. Was ist aber stattdessen das einheitssichernde Kriterium? Parfit vertritt einen an John Locke anknüpfenden psychologischen Reduktionismus, dem zufolge bestimmte psychologische Tatsachen und deren angemessene diachrone Verbindung konstitutiv für die personale Identität sind.<sup>11</sup> Psychologische Tatsachen zeichnen sich dadurch aus, dass sie beschrieben werden können, ohne einen eigenständigen Träger vorauszusetzen oder ausdrücklich zu behaupten, dass die Erfahrungen im Leben dieser Personen vom »Selbst« gemacht werden.<sup>12</sup>

Die verschiedenen Formen der relevanten Verbindungen bestimmt Parfit nun wie folgt: Zum einen muss zwei zeitlich aufeinanderfolgenden Bewusstseinszuständen *X* und *Y* derselbe psychologische Gehalt zukommen, zum anderen müssen sie eine gewisse Kontinuität aufweisen. Parfits Überlegung ist hier, dass diese Kontinuität genau dann vorliegt, wenn es eine ausreichend große Anzahl von direkten psychologischen Verbindungen gibt und diese Verbindungen überlappende Ketten bilden. So sind etwa meine jetzigen psychologischen Zustände aus den vorhergehenden entstanden und diese wiederum teilweise aus vorhergehenden. Man kann daher überlappende Ketten solcher Verbindungen ausmachen, die zeigen, dass ich mit der Person identisch bin, die ich heute Morgen war, obwohl möglicherweise nur sehr wenige direkte Verbindungen zwischen meinen jetzigen psychischen Zuständen und denen von heute Morgen bestehen.<sup>13</sup>

11 Einen solchen Reduktionismus beschreibt Locke etwa in der folgenden Passage: »Wenn wir etwas sehen, hören, riechen, schmecken, fühlen, überlegen oder wollen, so wissen wir, daß wir das tun. Das gilt jederzeit hinsichtlich unserer gegenwärtigen Situation und Wahrnehmung; jeder wird dadurch für sich selbst zu dem, was er sein eigenes Ich nennt. [...] Hierbei kommt es in diesem Fall nicht darauf an, ob dasselbe Selbst in derselben oder in verschiedenen Substanzen weiterbesteht. Denn da das Bewußtsein das Denken stets begleitet und jeden zu dem macht, was er sein Selbst nennt und wodurch er sich von allen anderen denkenden Wesen unterscheidet, so besteht hierin allein die Identität der Person, das heißt das Sich-selbst-gleich-Bleiben eines vernünftigen Wesens.« (Locke 1981 [1689], S. 419 f.) Für einen Überblick über Lockes Theorie der personalen Identität vgl. Quante 2007, Kap. 3.

12 Vgl. RaP, S. 210.

13 Vgl. ebd., S. 206.

Das allein ist sicherlich schon eine recht diskutabile Position, die den »Sturm« der Parfit-Rezeption erklären könnte.<sup>14</sup> Ein weiterer Grund für Parfits Prominenz in der Debatte ist aber auch darin zu sehen, dass er sich nicht mit der Entwicklung einer metaphysisch-ontologischen Position begnügt, sondern, wie oben angedeutet, zugleich auch lebenspraktische *normative* Konsequenzen daraus ableitet. Eine solche betrifft etwa unseren Umgang mit dem Tod, der sich durch eine reduktionistische Identitätskonzeption verändern würde. Das macht Parfit an einer Variation des Teletransporter-Beispiels deutlich.<sup>15</sup> In dieser kommt es bei der Replikation zu einer Panne, sodass die ursprüngliche Person nach dem Kopiervorgang nicht zerstört wird. Es gibt also zweimal dieselbe Person, eine auf der Erde und eine auf dem Mars. Da nun zwei Personen in psychologischer Kontinuität mit der früheren Person stehen, könne keine eindeutige Identitätsbeziehung mehr angegeben werden; dies sei aber kein Manko, da wir – so die Bedeutungslosigkeitsthese – dennoch alle relevanten Tatsachen über die Welt kennen. Allerdings ist, so das Beispiel weiter, das Herzkreislaufsystem des Originals beim Kopieren beschädigt worden, so dass diese Person in den folgenden Tagen sterben wird. Das aber sollte diese laut Parfit nicht sonderlich beunruhigen. Sie sollte vielmehr gelassen bleiben, denn auf dem Mars wird es eine Fortsetzung ihres Lebens geben (wenn auch nicht sie es leben wird und es womöglich noch nicht einmal dasjenige Leben *ist*, das sie bisher gelebt hat). Es ist ja nach Parfit die psychologische Kontinuität, welche das Weiterleben sichert, und diese Persistenz wird durch den Kopiervorgang gesichert. Entsprechend gilt: »Zerstört und repliziert zu werden ist in etwa so gut wie das gewöhnliche Überleben.«<sup>16</sup> Parfit zufolge impliziert die Bedeutungslosigkeitsthese also eine gewisse Gelassenheit gegenüber der eigenen Sterblichkeit.<sup>17</sup>

14 Manche – etwa Martina Herrmann – sehen in Parfit sogar den Initiator einer »quasi poststrukturalistische[n] Wende« (Herrmann 2001, S.179) in der analytischen Philosophie. Und dass aus dieser Deutungsperspektive heraus die Parfit'sche Theorie nicht gerade Jubelstürme ausgelöst hat, versteht sich angesichts der in der Philosophie vorherrschenden Skepsis gegenüber dem Poststrukturalismus beinahe von selbst.

15 RaP, S. 200.

16 Ebd., S. 201 (Übers.: Hg.).

17 »Als ich annahm, [dass die personale Identität eine weitere wichtige Tatsache darstellte,] war ich auch mehr über meinen eigenen unausweichlichen Tod be-

Eine weitere normative Konsequenz sieht er in der Einstellung gegenüber den Mitmenschen. In einer berühmt gewordenen Passage schreibt er über die Implikationen der Bedeutungslosigkeitsthe-  
these:

Ist die Wahrheit bedrückend? Einige dürften das so sehen. Ich aber finde sie befreiend und tröstlich. Als ich annahm, dass meine Existenz eine eigenständige Tatsache darstellte, schien ich in mir selbst gefangen. Meine Leben schien mir wie ein Glastunnel, durch den ich mich jedes Jahr schneller fortbewegte und an dessen Ende Dunkelheit wartete. Seit ich meine Sichtweise geändert habe, sind die Wände meines Glastunnels verschwunden. Ich lebe jetzt unter freiem Himmel. Es besteht immer noch ein Unterschied zwischen meinem Leben und dem anderer Menschen. Aber dieser Unterschied ist geringer geworden. Mein mir noch verbleibendes Leben kümmert mich nun weniger, das Leben der anderen dagegen mehr.<sup>18</sup>

Im Ergebnis läuft Parfits Slogan von der Bedeutungslosigkeit unserer Identität also auf eine Einstellungs- und Verhaltensänderung hinaus, die unsere eigene Existenz und die Beziehung zu anderen betrifft – und diese Neuorientierung hält er, im Sinne der normativ-ethischen Lesart von »what matters«, für einen Fortschritt, den es gemäß seiner normativen Theorie gutzuheißen gilt: Es kann vernünftig sein, das Wohlergehen anderer dem eigenen Vorteil vorzuziehen (siehe insbesondere Abschnitt 2.4).

sorgt. Nach meinem Tod wird es kein lebendes Wesen geben, das ich sein werde. Jetzt kann ich diese Tatsache reformulieren. Obwohl es auch später viele Erfahrungen geben wird, wird keine dieser Erfahrungen über Ketten direkter Verbindungen mit meinen momentanen Erfahrungen verknüpft sein, wie sie in meinem Erlebnisgedächtnis bestehen. [...] Nachdem ich das nun erkannt habe, erscheint mir mein Tod weniger schlimm.« (RaP, S. 281; Übers.: Hg.)

18 Ebd. (Übers.: Hg.).

### 1.3 Die Rezeptionsgeschichte der Theorie personaler Identität

Schon vor der Veröffentlichung von RaP wurden Parfits Thesen zur diachronen Identität breit und kontrovers diskutiert. Mit der Buchveröffentlichung von 1984 gewann die Diskussion aber stark an Lebhaftigkeit und thematischer Breite. Es wurden mitunter sehr detaillierte Auseinandersetzungen über einzelne Theoriebausteine geführt, und die Rezeptionsgeschichte ist entsprechend umfangreich.<sup>19</sup> Wir beschränken uns daher auf die Darstellung einiger ausgewählter Repliken: Zum einen auf solche, die die internen Verflechtungen der Theoriebausteine in Parfits Denken deutlich machen, zum anderen auf solche, die Anlass für eine konstruktive Weiterarbeit gegeben haben – sei es für Parfit selbst, sei es für andere Autoren.

In methodischer Hinsicht haben einige Kritiker bereits früh auf eine Engführung in der von Parfit präsentierten Grundkonstellation hingewiesen. So haben Paul Snowdon und Eric Olson angemerkt, dass Parfit nur unzureichend auf den Animalismus eingeht, welcher den Organismusbegriff in den Mittelpunkt stellt und den Begriff der biologischen Kontinuität als Kriterium in Anschlag bringt.<sup>20</sup> Zu dieser neueren Alternative hat sich Parfit erst in der jüngeren Vergangenheit geäußert, insbesondere in dem hier vorliegenden Aufsatz »We are not Human Beings« von 2012. Eignigermaßen überraschend ist hierbei allerdings seine Konklusion: Parfit verteidigt dort nämlich nicht seine ursprüngliche lockeanische Position, sondern präsentiert eine gänzlich neue Variante: »Einer anderen, besseren Sichtweise zufolge sind wir keine Tiere oder menschlichen Lebewesen. Wir sind das, was McMahan die bewussten, denkenden und kontrollierenden *Teile* menschlicher Lebewesen nennt. Wir können dies die *Verkörperter-Teil-Position* nennen.«<sup>21</sup> Parfit schlägt mithin nun doch eine substantielle En-

19 Siehe für die angelsächsische Diskussion etwa die Beiträge in Dancy 1997. Für die deutschsprachige Diskussion zum Personenbegriff im Allgemeinen und Parfit im Besonderen siehe etwa die Beiträge in Sturma 2001 sowie die kritischen Auseinandersetzungen mit Parfits Konzeption in Sturma 1997, VIII. 3 und Quante 2002, Kap. 5.2.1.

20 Vgl. Olson 1997 und 2002 sowie Snowdon 1996.

21 In diesem Band, S. 117.

tität als zentrales Element für personale Identität vor und nennt sie auch beim Namen: »All die für uns charakteristische mentale Aktivität hängt von unserem Zerebrum ab.«<sup>22</sup> Freilich meint Parfit, dass es sich nur um einen kleinen »Kunstgriff« handelt, der seine grundsätzlich lockeanische Position (samt ihren lebenspraktischen Konsequenzen) nicht in Mitleidenschaft zieht. Ob er sich damit aber nicht andere Herausforderungen aufbürdet, die gegebenenfalls noch schwieriger zu meistern sind, und ob das Verhältnis zu den anderen Bausteinen seiner Theorie nicht doch problematisch sein könnte, ist Gegenstand der gegenwärtigen Forschung.<sup>23</sup>

Darüber hinaus wurden auch eine Reihe von metatheoretischen Positionen diskutiert. Besonderes Interesse erregte hierbei immer wieder Parfits Verknüpfung von metaphysischen und ethischen Thesen. Einer einschlägigen Entgegnung Ernest Sosas nach ist es eine Sache, etwas über die transtemporalen Einheitsbedingungen der Person zu wissen; etwas anderes hingegen, wie man sich zu diesen metaphysischen Einsichten verhalten soll: »Die Logik allein wird nicht bestimmen, wie wir uns entscheiden.«<sup>24</sup> Das ist sicherlich erst einmal nur eine Behauptung. Es zeigt aber, dass noch nicht alle Details von Parfits Theorie der diachronen Identität geklärt sind. Entweder wäre an dieser Stelle eine metaphilosophische Absicherung vonnöten, die die direkte Verbindung von Metaphysik und Ethik begründet – oder es wäre zu rechtfertigen, warum die von ihm lancierte Einstellungsänderung für gut gehalten werden sollte.

Hierbei handelt es sich in jedem Fall um eine normative Frage, die ausschließlich im Rahmen einer ethischen Theorie zu beantworten ist. Nun hat Parfit in Teilen von RaP gerade eine solche ethische Theorie vorgelegt. Aber reicht das aus? Das haben die Kritiker wiederum bestritten. So hat Peter Unger darauf hingewiesen, dass die veränderte Einstellung möglicherweise einige Vorteile bringt – die Linderung von Todesangst und die Abkehr vom Egoismus –, zugleich aber einige nicht zu vernachlässigende Kosten aufweist, welche er unter dem Titel »Fokusverlust« [*loss of focus*] in den Kapiteln sieben bis neun seines Buchs *Identity, Consciousness and Value* (1990) abhandelt. Aber wie auch immer man diese Abwägung vornimmt, deutlich geworden ist in der Diskussion, dass die Über-

22 In diesem Band, S. 112.

23 Vgl. exemplarisch für direkte Repliken Olson 2015, Bailey 2016 und McGee 2016.

24 Sosa 2003, S. 201 (Übers.: Hg.).

zeugungskraft der normativen Implikationen keineswegs nur von metaphysischen Überlegungen gespeist wird, sondern ihrerseits in die normativ-ethische Theorienbildung weist.

Eine weitere zentrale These Parfits, die immer wieder in die Kritik geraten ist, betrifft die Behauptung, dass die psychologische Kontinuität hinreichend für die Zuschreibung von diachroner Identität ist. Hierbei handelt es sich sicherlich um den zentralen »Fluchtpunkt«, von dem ausgehend Parfit seine Theorie der Person entwickelt. Entsprechend umfangreich ist auch die Rezeption ausgefallen. In ihrer Struktur sind die vielen Kritiklinien aber weitestgehend einheitlich: Es geht immer darum, auf die eine oder andere Weise die Kontraintuitivität von Parfits lockeanischer Theorie herauszuarbeiten.

Die Beispiele und Gedankenexperimente, die das zeigen sollen, sind recht heterogen. Greifen wir ein Beispiel heraus, das in der Rezeption immer wieder aufgenommen wurde und auch für die weitere Theoriebildung Parfits von Bedeutung war, das insbesondere von Mark Johnston vorgestellte Beispiel der *singular goods*.<sup>25</sup> Hierbei handelt es sich um Arten von Projekten, bei denen es nicht nur darauf ankommt, dass sie realisiert werden, sondern auch, dass eine bestimmte Person dies tut. Wer seinen Traumpartner heiraten möchte, dem ist nicht nur daran gelegen, dass der Traumpartner jemanden heiratet, sondern dass er selbst dieser jemand ist. Das kann Parfit jedoch nicht zugestehen. Für ihn ist die psychologische Kontinuität hinreichend, um die diachrone Identität zuzuschreiben. Es reicht daher aus, dass der Heiratswillige mit dem Individuum, das tatsächlich geheiratet wird, in großen Teilen die gleichen psychologischen Tatsachen aufweist. Das führt aber zur kontraintuitiven Ansicht, dass es *singular goods* nicht gibt.

Sollten wir bei solchen Implikationen nicht eher die Theorie aufgeben, als unsere alltagspraktisch verankerte Intuition zu revidieren? Diese Frage würde Johnston sicherlich bejahen, Parfit hingegen würde, wie wir an seinen Beispielen ablesen können, mit dieser Kontraintuitivität kaum ein Problem haben. Gleichzeitig muss man ihm zugutehalten, dass er weitere Überlegungen aus der Rationalitätstheorie anführt, um seine Ansicht zu untermauern. So versucht er gerade im zweiten Teil von RaP, unabhängig von sei-

25 Vgl. Johnston 2003.

ner Theorie personaler Identität nachzuweisen, dass alle rationalen Einstellungen aus einer unpersönlichen Perspektive stammen, die keine Referenz auf das Subjekt enthalten. Insofern beruht die Bezugnahme auf *singular goods* laut Parfit nicht nur auf einer falschen Vorstellung von diachroner Identität, sondern ist zusätzlich noch schlicht arational, da es sich nicht um unpersönliche Gründe handelt. Allerdings versteht sich auch dieser Rationalitätstheoretische Theoriebaustein keineswegs von selbst, was sich an der mitunter harschen Kritik gezeigt hat.<sup>26</sup> Parfit selbst hat auf diese Repliken mittlerweile reagiert und in OWM einen neuen Versuch unternommen, seine Theorie praktischer Gründe weiter zu explizieren und damit zumindest implizit die offene Flanke zu verteidigen.

## 2. Moralphilosophie

Ein grundlegender Impuls, der Parfit bei der Behandlung moralphilosophischer Fragen leitet, ist die Überzeugung, dass zwischen Moralphilosophen weitaus mehr Einigkeit möglich sein muss, als der Blick auf die klassischen Streitlinien der konkurrierenden Theorieansätze und die damit verbundene »Lagerbildung« vermuten lässt. Dahinter steht Parfits Überzeugung, dass Meinungsverschiedenheiten zwischen Philosophen, die sich nicht auf Missverständnisse oder auf fehlende Informationen zurückführen lassen, Belege für die These darstellen können, dass es letztlich gar keine Wahrheiten über normative Sachverhalte gebe.<sup>27</sup> Dann aber wäre möglicherweise gar nichts von Bedeutung und die Moral eine bloße Illusion:

Wenn wir unsere Meinungsverschiedenheiten nicht beilegen können, würde uns das Gründe geben zu bezweifeln, dass es überhaupt wahre Prinzipien gibt. Es könnte sein, dass es gar keine wirkliche Moral gibt – sie wäre lediglich eine Illusion.<sup>28</sup>

26 Siehe hierfür unter anderem die Rezeptionsgeschichte seiner Theorie praktischer Gründe in Abschnitt 3.6.

27 Dass Parfits Moralphilosophie durchgängig als Auseinandersetzung mit dieser drohenden Konsequenz gelesen werden muss, betont etwa Scheffler 2011.

28 OWM, Bd. 2, S. 151 (Übers.: Hg.).

Vor dem Hintergrund seines Vermittlungsgedankens überrascht es nicht, dass Parfit zwei äußerst gegensätzliche Denker als geistige Ziehväter anführt:<sup>29</sup> Auf der einen Seite den wichtigsten Vertreter der deontologischen Ethik, Immanuel Kant (1724-1804), auf der anderen Seite den britischen Utilitaristen Henry Sidgwick (1838-1900).

Kant – den er erst im Zuge einer systematischen Ausarbeitung seiner normativen Ethik seit den 1990er Jahren für sich entdeckt hat – gilt Parfit vor allem als kreativer Kopf, dessen genialische Ideen reichlich Anregungen zum kritischen Weiterdenken geben. Zugleich seien dessen Texte aber von einer erstaunlichen Inkonsistenz geprägt. So bewundert er einerseits Kants *Grundlegung zur Metaphysik der Sitten* für ihren unglaublich produktiven Gehalt («auf gerade mal vierzig Seiten brennt Kant ein Feuerwerk ab, das uns mehr neue und fruchtbare Ideen liefert als alle Philosophen der vorangegangenen Jahrhunderte zusammen»), hält deren Grundthesen aber gleichzeitig für widersprüchlich: »Zu all den Qualitäten, die Kant so viel erreichen lassen, zählt auch seine Inkonsistenz.«<sup>30</sup> Ein Teil von Parfits normativer Ethik besteht daher in dem Versuch, die verschiedenen Ideen Kants zu systematisieren, durch verschiedene Reformulierungen exakt definierte Moralprinzipien voneinander zu unterscheiden und schließlich zu evaluieren, welche dieser Prinzipien eine tragfähige Theorie abgeben könnten.

An Sidgwick interessieren Parfit einerseits dessen begriffsanalytische Leistungen für die konsequentialistische Traditionslinie, stärker aber noch die *Methode* der moralphilosophischen Argumentation. Sidgwick habe zwar nicht den genialischen Charakter Kants, arbeite dafür aber exakt und vertrete deshalb im Vergleich zu seinen Vorläufern die weitaus größere Zahl wahrer und bedeutender Thesen.<sup>31</sup> Parfit bezeichnet Sidgwicks *Methods of Ethics* daher als »das beste moralphilosophische Buch, das je geschrieben wurde.«<sup>32</sup> Wie Allen Wood treffend feststellt, scheint Parfit insbesondere an Sidgwicks Herangehensweise interessiert zu sein, Moralprinzipien im Lichte der alltäglichen moralischen Überzeugungen der Menschheit aufzustellen und zu testen. Sidgwick geht nämlich bewusst

29 Vgl. OWM, Bd. 1, S. xxxiii.

30 Ebd., S. 183 (Übers.: Hg.).

31 Vgl. ebd., S. xxxiii.

32 Ebd. (Übers.: Hg.).

nicht von übergeordneten Überlegungen aus, von denen moralische Urteile dann abgeleitet werden, sondern übernimmt Moralprinzipien aus dem Common Sense, testet sie auf Konsistenz und sucht nach Beispielen, bei denen die Moralprinzipien zu Ergebnissen führen, die unserer Intuition massiv widersprechen.<sup>33</sup>

Um unsere Alltagsintuitionen für die moralphilosophische Argumentation heranziehen zu können, bezieht sich Parfit immer wieder auf Gedankenexperimente. Dies hat zum Ziel, jeweils möglichst wenige moralisch bedeutsame Eigenschaften von Handlungen einzeln sichtbar und damit intuitiv bewertbar zu machen. Gerade aufgrund dieser massiven Reduktion auf wenige Aspekte erscheinen die Fallbeispiele in der Regel als äußerst unrealistisch und weit hergeholt. Moralprinzipien, die allgemeine Gültigkeit beanspruchen, müssen laut Parfit aber auch dann zu plausiblen Ergebnissen führen, wenn man sie auf solche Beispiele anwendet.

## 2.1 Das Problem der Nicht-Identität und unsere Pflichten gegenüber kommenden Generationen

Die Auseinandersetzung Parfits mit Themen der Moralphilosophie im engeren Sinn beginnt mit einer einfachen Frage: Haben wir eine moralische Pflicht, so zu handeln, dass es kommenden Generationen möglichst gut geht? Bei näherem Hinsehen führt diese Frage auf ein schwieriges Problem, das Parfit bereits 1976 andeutet und dann in RaP äußerst wirkmächtig formuliert.<sup>34</sup>

Entscheiden wir uns für eine Maßnahme, die das Wohlergehen künftiger Generationen befördern soll, so hat dies auch Auswirkungen auf unser eigenes Leben. Setzen wir etwa Regeln in Kraft, die den CO<sub>2</sub>-Ausstoß von Autos reduzieren sollen, so werden die meisten Menschen vielleicht seltener zur Tankstelle fahren, und ihr Tagesablauf wird sich dadurch geringfügig ändern. Dadurch werden sich aber auch die genauen Zeitpunkte und Umstände verschieben, zu denen Menschen Geschlechtsverkehr haben, und infolgedessen werden sich andere Spermien bei einer möglichen Befruchtung durchsetzen. Das bedeutet, dass banale Änderungen

33 Vgl. Wood 2011, S. 59.

34 Vgl. aber auch die älteren Texte Parfit 1976 und Parfit 1982. Parfit diskutiert das Problem erneut in OWM, Bd. 2, S. 217 ff., im Zusammenhang mit einer Kritik an Scanlons Kontraktualismus.

unserer Lebensgewohnheiten, die durch eine politische Maßnahme hervorgerufen werden, dazu führen, dass innerhalb weniger Generationen ganz andere Menschen existieren als diejenigen, die existiert hätten, wenn wir die fragliche Maßnahme nicht umgesetzt hätten.

Daraus ergibt sich folgendes Problem: Angenommen, wir stehen vor der Frage, ob wir mit Klimaschutzmaßnahmen die Lebensqualität künftiger Generationen verbessern wollen. Entscheiden wir uns dagegen, können künftige Generationen vielleicht nur noch ein sehr eingeschränktes (wenngleich immer noch lebenswertes) Leben führen; entscheiden wir uns dafür, haben sie annähernd unsere Möglichkeiten der Lebensgestaltung. Intuitiv scheint klar zu sein, dass wir moralisch verpflichtet sind, die Klimaschutzmaßnahmen umzusetzen. Jedoch gilt aufgrund des oben Gesagten: Wenn wir uns für die Maßnahme entscheiden, werden ganz andere Menschen existieren als in dem Fall, in dem wir uns gegen die Maßnahme entscheiden. Gegenüber den Menschen, die existieren werden, wird deshalb mit keiner Handlungsoption eine Pflicht verletzt, denn deren Wohlergehen wird – soweit das von uns abhängt – absurderweise in jedem Fall durch unsere Handlung maximiert, denn mit jeder Entscheidung für eine andere Handlungsoption würden sie gar nicht existieren. Daher stellt sich die Frage, aus welchem Grund es falsch wäre, die Klimaschutzmaßnahme zu unterlassen, wenn wir doch dadurch niemanden schlechter stellen: »Da [unsere Entscheidung] niemanden schlechter stellt, müssen wir erklären, warum wir einen moralischen Grund haben, [diese Entscheidung] nicht zu treffen.«<sup>35</sup>

Mit einer Theorie, die die Moral allein auf Rechte von Personen zurückführt, lasse sich das Problem der Nicht-Identität prinzipiell nicht lösen. Zugleich sei die Intuition, dass wir moralisch verpflichtet sind, künftigen Generationen ein möglichst hohes Wohlergehensniveau zu ermöglichen, nicht zurückweisbar. Parfit zieht daraus den Schluss, dass Moraltheorien, die rein auf Rechten basieren, abzulehnen seien. Stattdessen müsse für jede plausible Moraltheorie ein Zustand, in dem es Menschen gut geht, in einem nicht personenbezogenen Sinn für besser gehalten werden als ein Zustand, in dem es der gleichen Zahl von Menschen schlecht geht.

35 RaP, S. 378 (Übers.: Hg.).

Eine plausible Moraltheorie muss also behaupten, dass wir einen unparteiischen Grund haben zu wünschen und zu bewirken, dass es viele Menschen gibt, denen es möglichst gut geht. Damit ist bereits vorgezeichnet, dass Parfits Moralphilosophie erstens von der Grundannahme lebt, dass es normative Gründe *unabhängig von den subjektiven Einstellungen existierender Menschen* geben muss – Parfit spricht davon, dass normative Gründe *objektgegeben* sind (siehe auch Abschnitt 3.2); und dass zweitens nicht alle normativen Gründe auf das Eigeninteresse zurückgeführt werden, sondern auch in einer unparteiischen Perspektive bestehen können.

Eine Theorie, die der Existenz von objektgegebenen, unparteiischen Gründen Rechnung trägt, gerät allerdings in ein anderes Problem: Sie impliziert womöglich die sogenannte Abstoßende Schlussfolgerung.

## 2.2 Populationsethik: Die Abstoßende Schlussfolgerung

Um das Problem der Nicht-Identität zu umgehen, müssen verschiedene mögliche Zustände der Welt, in denen jeweils unterschiedliche Menschen leben, aus moralischer Sicht miteinander vergleichbar sein. Als Bewertungskriterien, die einen möglichen Zustand besser machen als einen anderen, kommen dabei sowohl die Qualität (insbesondere das Wohlergehen) als auch die Quantität der einzelnen Leben in Frage: Wenn es beispielsweise einer konstanten Zahl an Menschen *besser* geht, macht dies einen Zustand besser; und wenn es *mehr* Menschen gibt, die ein lebenswertes Leben führen (deren Leben also alles in allem mehr Gutes als Schlechtes enthält), macht dies einen Zustand ebenfalls besser.

Stellen wir uns zur Abgrenzung kurz eine Theorie vor, der zufolge die Zahl der Menschen *keine* Rolle für die Bewertung eines Zustands spielt: Innerhalb einer solchen Theorie gäbe es keinen Grund zu wünschen, dass eher eine Welt mit 100 stark leidenden Menschen existiert, als dass eine Welt mit 10 Millionen stark leidenden Menschen existiert. Eine solche Theorie wäre dem zusätzlichen Leid gegenüber unsensibel und sollte daher verworfen werden.

Wenn nun aber andererseits sowohl das Wohlergehen als auch die Anzahl der Menschen eine Rolle spielt, kann jedwede Verminderung an Wohlergehen durch eine hinreichend große Zahl von Menschen ausgeglichen werden. Eine Welt, in der es eine Million

Menschen gibt, denen es sehr gut geht – beziffern wir ihr Wohlergehensniveau schematisch mit 100 –, wäre dann weniger gut als eine Welt, in der es eine *Milliarde* Menschen gibt, denen es fast genauso gut geht, deren Wohlergehen also etwa bei 99 oder 98 liegt. Diese Argumentation lässt sich aber immer weiterführen – so weit, bis es Menschen nur noch gerade eben so gut geht, dass die Gesamtbilanz ihres Lebens positiv ausfällt und ihr Leben gerade eben noch lebenswert ist. Für jede beliebige Zahl von Menschen, denen es sehr gut geht, gilt daher, dass es besser wäre, wenn eine viel größere Zahl von Menschen existierte, die gerade eben noch lebenswerte Leben führen. Während die Prämissen der Argumentation vielen einleuchten, wehren sich die meisten Menschen gegen die Konklusion, die Parfit als »Abstoßende Schlussfolgerung« bezeichnet: Eine überbevölkerte Welt, in der Menschen gerade noch lebenswerte Leben führen, kann unmöglich besser sein als eine Welt, in der eine immerhin beträchtliche Anzahl von Menschen äußerst wertvolle Dinge verwirklicht und einen sehr hohen Lebensstandard genießt.

Parfit hat diese Überlegung als philosophisches Problem in die ethische Debatte eingeführt und damit eine enorme Diskussion ausgelöst, die mittlerweile einen eigenen Zweig der Angewandten Ethik ausmacht: die Populationsethik (siehe Abschnitt 2.5). Wenngleich Parfit beständig nach einer Lösung des Problems sucht, wie an den Prämissen festgehalten werden kann, ohne dass die unliebsame Konklusion zugestanden werden muss – was ihm angesichts einer drohenden Inkonsistenz der Moraltheorie als äußerst dringlich erscheint –, ist er lange Zeit möglichen Lösungsvorschlägen gegenüber eher kritisch eingestellt. In dem 1986 erstmals erschienenen Aufsatz »Überbevölkerung und Lebensqualität«, der im vorliegenden Band enthalten ist, fasst er zunächst die Problembeschreibung aus RaP zusammen und skizziert dann einen perfektionistischen Lösungsansatz: Möglicherweise wäre der komplette Wegfall der besten Dinge, die das Leben lebenswert machen (Parfit denkt hier beispielsweise an Mozarts Musik), durch keine Zahl von Menschen zu kompensieren, die über diese besten Dinge gar nicht mehr verfügten. Allerdings betont Parfit auch, dass sich der Perfektionismus selbst einigen Einwänden ausgesetzt sieht.<sup>36</sup>

Erst in seiner letzten Lebensphase tritt er offener für eine wei-

<sup>36</sup> Vgl. auch Parfit 2004a.

ter entwickelte Variante dieser Lösung ein. 2016 publiziert Parfit unter dem Titel »Can We Avoid the Repugnant Conclusion?« die Druckfassung eines Vortrags, die im vorliegenden Band ebenfalls enthalten ist. Parfit möchte in diesem Text den Perfektionismus in eine Position einbetten, die er als »Ungenau Lexikalische Sichtweise« bezeichnet. Sein Ausgangspunkt ist die von der Philosophin Ruth Chang entwickelte These, dass beim Vergleich zweier Zustände einer nicht nur besser oder schlechter sein kann als der andere bzw. beide gleich gut sind, sondern dass eine *vierte* Möglichkeit besteht, der zufolge beide einander »ebenbürtig« [*on a par*] sein können.<sup>37</sup> Changs Beispiel ist die Genialität von Personen, die in ganz verschiedenen künstlerischen Disziplinen tätig sind, etwa Michelangelo und Mozart. Während beide klarerweise besser sind als mittelmäßige bildende Künstler bzw. Musiker, erscheint es nicht sehr vielversprechend, sie im direkten Vergleich als »gleich gut« einzustufen. Besser sei es, sie in künstlerischer Hinsicht als einander ebenbürtig aufzufassen.

Parfit interpretiert »ebenbürtig« nun als »in ungenauer Weise gleich gut« und argumentiert, dass sich auch Wohlergehenszustände aufgrund des qualitativen Unterschieds der in ihnen realisierten Güter nur *ungenau* miteinander vergleichen lassen und also in ungenauer Weise besser, schlechter oder gleich gut sein können.

Die Pointe dieser Argumentation ist, dass »in ungenauer Weise gleich gut« keine transitive Relation mehr darstellt. Damit ist gemeint: Ist die Welt *A* in ungenauer Weise so gut wie *B* und die Welt *C* in ungenauer Weise so gut wie *B*, dann ist noch lange nicht gesagt, dass *C* in ungenauer Weise so gut wie *A* ist. *C* könnte auch in ungenauer Weise besser oder schlechter sein. Denken wir an das Beispiel, in dem ein Maler mit einem Musiker verglichen wird, diesmal aber mit drei Personen: Es ist denkbar, dass Maler *A* in ungenauer Weise gleich gut ist wie Musiker *B* und dass ebenso Maler *C* in ungenauer Weise gleich gut ist wie Musiker *B*. Es bleibt in diesem Fall möglich, dass Maler *A* und Maler *C* aufgrund der besseren Vergleichbarkeit nicht als in ungenauer Weise gleich gut bewertet werden, sondern dass einer von beiden als besser einzustufen ist.

In gleicher Weise kann man deshalb über den Vergleich von Weltzuständen behaupten, dass gilt:

37 Vgl. Chang 1997; 2002 und 2016.

Eine Welt mit 10 Millionen Menschen auf Niveau 100 ist in ungenauer Weise gleich gut wie eine Welt mit 15 Millionen Menschen auf Niveau 95, und diese ist wiederum in ungenauer Weise gleich gut wie eine Welt mit 30 Millionen Menschen auf Niveau 90 und so fort bis hin zu einer Welt mit 300 Milliarden Menschen auf Niveau 2,

ohne dass man dadurch logisch darauf festgelegt wäre zu behaupten, dass gilt:

Eine Welt mit 10 Millionen Menschen auf Niveau 100 ist in ungenauer Weise gleich gut wie eine Welt mit 300 Milliarden Menschen auf Niveau 2.

Daher kann ohne Widerspruch behauptet werden, dass eine Welt mit hinreichend vielen Menschen auf Niveau 100 *immer besser ist* als eine Welt mit *noch so vielen* Menschen auf Niveau 1. Wie Parfits weitere Überlegungen aber zeigen, lässt sich auch unter dieser Prämisse die Abstoßende Schlussfolgerung nur mithilfe perfektionistischer Prämissen erfolgreich vermeiden.<sup>38</sup>

### 2.3 Was ist schlecht an Ungleichheit? Der Prioritarismus

Parfits Moralphilosophie prägt noch ein weiteres Feld innerhalb der übergeordneten Fragestellung, wie Zustände oder mögliche Welten miteinander verglichen werden sollten: Wie sind Zustände zu bewerten, in denen das Wohlergehen der Menschen ungleich verteilt ist? Macht Ungleichheit einen Zustand weniger wünschenswert, und wenn ja, warum? Insbesondere an die Diskussion um John Rawls' *A Theory of Justice* anknüpfend, bildete der Egalitarismus in den 1980er und 1990er Jahren die dominante Strömung der politischen Philosophie, aber auch darüber hinaus in der normativen Ethik. Kontrovers war vor allem, *in Bezug auf was* Gleichheit herzustellen sei: Chancen, Grundgüter, Ressourcen, Wohlergehen und weitere Möglichkeiten kommen hier in Frage.<sup>39</sup> Dass Gleichheit

38 Vgl. die Auseinandersetzung mit dem *Auf-und-ab-Argument*, in diesem Band, S. 420.

39 Rawls (1971) steht paradigmatisch für die Idee gleicher Freiheitsrechte und Chancen zur Erlangung bevorzugter Positionen; Dworkin (1991) zielt wirk-

den Grundbegriff von Gerechtigkeitstheorien bilden sollte, war dagegen weitgehend unstrittig.<sup>40</sup> Diese Hintergrundannahme stellt Parfit 1991 im Rahmen der *Lindley Lecture* in einem Vortrag in Frage (veröffentlicht wurde der Text zunächst 1995 als eigenständige Broschüre unter dem Titel *Equality or Priority?*;<sup>41</sup> im vorliegenden Band ist die Übersetzung der zusammenfassenden Version »Equality and Priority« von 1997 abgedruckt<sup>42</sup>). Einerseits legt Parfit darin offen, dass Gleichheit aus ganz verschiedenen Gründen wertgeschätzt werden kann. Andererseits zeigt er, dass viele unserer moralischen Überzeugungen entgegen der Standardauffassung nicht auf ein Gleichheitskonzept zurückzuführen sind.

Ungleichheiten zwischen Menschen können das Resultat von gerechtfertigten oder unfairen Verteilungsprozessen sein und daher evaluiert werden, indem die Verteilungsprozesse einer Bewertung unterzogen werden. Wir können Zustände der Ungleichverteilung aber auch isoliert von diesen Entstehungsprozessen bewerten. Das machen wir insbesondere dann, wenn wir der Meinung sind, dass die Verteilungsprozesse nicht von Menschen beeinflusst werden, etwa weil die Ungleichheit genetisch bedingt ist wie bei manchen Aspekten unserer Begabungen und körperlichen Fähigkeiten.

Ist auch in diesem Fall ein Zustand der Gleichverteilung in irgendeiner Hinsicht besser als ein Zustand der Ungleichverteilung? Wäre etwa eine Welt, in der alle über die gleichen körperlichen Fähigkeiten verfügen, besser als eine Welt, in der manche sehr gute Fähigkeiten haben, andere aber nicht? Nehmen wir als Beispiel den Zustand einer Welt, in der jeder ein Auge hat und also mäßig gut sehen kann, und vergleichen ihn mit dem Zustand einer Welt, in der die eine Hälfte der Bevölkerung zwei funktionierende Augen hat, die andere Hälfte aber blind ist. Um den philosophischen Kern des Problems in den Blick zu bekommen, ist es wichtig, von den Konsequenzen der Ungleichheit abzusehen. Tun wir also so, als gäbe es in der zweiten Welt keinen Neid der Blinden auf die Se-

mächtig auf Ressourcengleichheit; und Roemer (1998) verteidigt die Idee einer Chancengleichheit bei der Erlangung von Wohlergehen.

40 Vgl. die Belege in Krebs 2000, S. 11.

41 Der geläufige Wiederabdruck der umfangreichen Version ist Parfit 2002 [1995].

42 Vgl. Parfit 1997a. Die hier abgedruckte Übersetzung ist eine überarbeitete Fassung der Übersetzung, die in *Gleichheit oder Gerechtigkeit. Texte der neuen Egalitarismuskritik* (Krebs 2000) erschienen ist.

henden, weil die Personen nichts voneinander wissen; und nehmen wir an, dass beide Welten dieselbe Gesamtsumme an Wohlergehen erreichen, nur dass diese Summe unterschiedlich verteilt ist. Viele Menschen würden auch unter diesen Prämissen denken, dass die Welt der Einäugigen besser sei als die andere, und daraus folgern, dass Gleichverteilung selbst einen intrinsischen Wert hat, der diese Welt besser macht. Parfit nennt diese Position den Teleologischen Egalitarismus.

Wie Parfit zeigt, ist der Teleologische Egalitarismus dem *Einwand der Angleichung nach unten* ausgesetzt: Vergleichen wir eine Welt, in der die Hälfte zwei Augen hat und die andere Hälfte keine Augen, mit einer Welt, in der niemand Augen hat. Wer am intrinsischen Wert der Gleichheit festhält, muss zugestehen, dass diese zweite Welt zwar nicht alles in allem, aber doch *in einer Hinsicht* besser ist als die andere. Die Welt würde also besser, indem manche Menschen schlechter gestellt wären, es aber niemandem besser ginge. Das ist laut Parfit unplausibel.

Er schlägt daher vor, unsere Intuition, dass die Welt durch Gleichverteilung besser werden kann, anders zu erklären als durch den Teleologischen Egalitarismus. Laut der *Vorrangposition* bzw. laut dem *Prioritarismus* gilt, dass ein Zuwachs an Wohlergehen umso wichtiger ist, je schlechter es einer Person geht. Während im klassischen utilitaristischen Kalkül jeder Wohlergehenszuwachs einer bestimmten Quantität gleich viel zählt, ungeachtet der Frage, wer ihn bekommt, zählt im Prioritarismus ein Wohlergehenszuwachs unterschiedlich viel, je nachdem, wie gut es dem geht, der ihn bekommt: Kann das Wohlergehensniveau einer Person, die kaum ihre Grundbedürfnisse befriedigen kann, erhöht werden, so fällt das mehr ins Gewicht, als wenn das Wohlergehen eines Menschen, dem es bereits sehr gut geht, um den gleichen Wert erhöht würde. Diese Position kann erklären, weshalb wir eine Welt, in der alle ein Auge haben, einer Welt vorziehen, in der manche zwei und andere gar kein Auge haben, ohne auf die unplausible Aussage festgelegt zu sein, dass eine Angleichung nach unten in irgendeiner Hinsicht einen Gewinn darstellen würde.

Parfit hat damit ein doppeltes Ziel erreicht: Seine begriffliche Analyse hat *erstens* aufgedeckt, dass es vielen, denen eine Angleichung von Wohlergehensniveaus wichtig erscheint, möglicherweise gar nicht um Gleichheit als solche geht, sondern um die Vorrangig-

keit der Schlechtergestellten. *Zweitens* legen seine Gedankenexperimente nahe, dass der Prioritarismus eine aussichtsreiche Position darstellt, die besser mit einigen unserer alltäglichen Überzeugungen zusammenstimmt als ein Teleologischer Egalitarismus einerseits und ein purer Utilitarismus andererseits.<sup>43</sup>

## 2.4 Die Triple Theory

Auch wenn Parfits Überlegungen zur Abstoßenden Schlussfolgerung und zur Verteilungsgerechtigkeit typischerweise der Debatte um die beste Spielart des Konsequentialismus zugeordnet werden, weil sie nach der Aggregation des Wohlergehens vieler Personen fragen, ist ihre Bedeutung keinesfalls auf konsequentialistische Theorien beschränkt. Eine jede Theorie, die möglichen künftigen Zuständen *irgendeine* Bedeutung zuspricht, ist dem Aggregationsproblem ausgesetzt – und das dürfte auf fast alle Moralthorien zutreffen. Strittig unter den Anhängern verschiedener Moralthorien ist allerdings, *welche* Bedeutung mögliche künftige Zustände haben und wie die verschiedenen Gründe, die uns mögliche künftige Zustände sowie weitere moralisch relevante Tatsachen an die Hand geben, gegeneinander abzuwägen sind. Dabei steht unter anderem im Raum, wie Gründe des Eigeninteresses gegen unparteiische Gründe, von denen in den Abschnitten 2.1 bis 2.3 überwiegend die Rede war, abgewogen werden können.

Wir müssen daher nach einer übergeordneten Theorie suchen, die uns Auskunft darüber gibt, was aus moralischer Perspektive letztlich zu tun ist, wenn wir alle Gründe zusammennehmen. Bei der Suche nach einer solchen Theorie sehen wir uns allerdings besonders tiefgreifenden Konflikten ausgesetzt, die sogar daran zweifeln lassen, dass *überhaupt* irgendeine moralische Forderung objektive Geltung beanspruchen kann. Parfits anspruchsvollstes Ziel innerhalb der Moralphilosophie liegt deshalb – wie oben bereits vorweggenommen – in dem Beweis, dass sich mit der *Triple Theory* als einer solchen übergeordneten Theorie diese (scheinbaren) Konflikte so weit als möglich ausräumen lassen.

Die *Triple Theory* präsentiert Parfit erstmals im Rahmen der

43 In Parfit 2012 antwortet Parfit auf eine Reihe neuerer Kritiken, um die Plausibilität des Prioritarismus weiter zu untermauern.

*Tanner Lectures* von 2002;<sup>44</sup> Jahre später erscheint eine erheblich überarbeitete, auf unzählige Kommentare anderer Moralphilosophen reagierende Fassung als Teil II und Teil III des ersten Bandes von *On What Matters*. Parfits zentrale These ist, dass drei der vorherrschenden moralphilosophischen Theorien sich entgegen der gängigen Ansicht nicht widersprechen, sondern zu identischen Moralprinzipien führen. Kantianismus, Kontraktualismus und Konsequentialismus besteigen demnach denselben Berg nur von verschiedenen Seiten: Je weiter deren Vertreter in der Theoriebildung fortschreiten, desto näher kommen sie sich, bis sie sich auf dem Gipfel des Berges endlich treffen. Parfit diskutiert daher alle drei Theorien, wählt jeweils die seiner Ansicht nach plausibelste Variante aus und versucht zu zeigen, dass die verbliebenen drei Formen inhaltlich deckungsgleich sind, weil sie zu den gleichen Moralprinzipien führen.

Da sich Kontraktualisten ohnehin meist in der kantischen Tradition sehen, sieht sich Parfits Argumentation in dieser Hinsicht eher im Detail Hürden ausgesetzt.<sup>45</sup> Anders verhält sich der Versuch der Zusammenführung von Kantianismus und Konsequentialismus. Zwar gibt es immer wieder Versuche nachzuweisen, dass ein »Kantischer Utilitarismus«<sup>46</sup> denkbar sei oder dass Kant eigentlich eine Form des Konsequentialismus begründe.<sup>47</sup> Nach dem vorherrschenden Paradigma der Moralphilosophie der Gegenwart sind Kantianismus und Konsequentialismus aber einander entgegengesetzte, inkompatible Moralpositionen mit je eigenen Stärken und Schwächen. Parfits Theorie zielt darauf, diesen Gegensatz als scheinbaren auszuweisen.

Seine Argumentation umfasst dabei zwei Stufen: In Teil II von OWM diskutiert Parfit einzelne Prinzipien Kants, die man als Moralprinzipien mittlerer Ebene bezeichnen könnte, also Moralprinzipien, die für ihren jeweiligen Anwendungsbereich, aber nicht für jede Situation Geltung beanspruchen. Im Hauptteil der Argumentation – dessen Kernelemente in diesem Band unter dem

44 Parfit 2004b.

45 Vgl. insbesondere das Zugeständnis minimaler verbleibender Differenzen zwischen dem Kantischen und dem Scanlon'schen Kontraktualismus in OWM, Bd. 2, Kap. 23.

46 Vgl. Hare 1993.

47 Vgl. Cummiskey 1990 und 1996.

Titel »Ein Kantisches Argument für den Regelkonsequentialismus« abgedruckt sind – geht es (in Teil III von OWM) um Theorien, die ein oberstes Moralprinzip formulieren, anhand dessen Moralprinzipien mittlerer Ebene geprüft werden können und aus denen daher möglichst für alle moralphilosophischen Probleme eine Lösung ableitbar sein sollte. Auch wenn erst in Teil III die Vereinigung der drei konkurrierenden Moraltheorien erfolgt und Parfit betont, dass dieser Teil ohne Kenntnis von Teil II gelesen werden kann, übernimmt Teil II eine wichtige Funktion: Dort werden insbesondere diejenigen Prinzipien Kants diskutiert, die konsequentialistischen Überzeugungen den stärksten Widerstand leisten. Indem Parfit nachweist, dass manche Lesarten dieser Formeln zu äußerst unplausiblen Resultaten führen, andere Lesarten die anti-konsequentialistischen Schlussfolgerungen jedoch gar nicht zulassen, die aus ihnen traditionell gezogen werden, bereitet er seine These vor, dass aus dem Kantianismus regelkonsequentialistische Prinzipien abgeleitet werden können. Auch wenn dies nicht das einzige Anliegen ist, das Parfit in Teil II verfolgt, sei im Folgenden beispielhaft eine Argumentation dieser Art vorgestellt.

Kants Selbstzweckformel zufolge sollen wir uns selbst und andere »jederzeit zugleich als Zweck, niemals bloß als Mittel«<sup>48</sup> behandeln. Prinzipiell stimmt Parfit zu, dass es immer falsch ist, andere als bloßes Mittel anzusehen.<sup>49</sup> Allerdings sehe in einer plausiblen Lesart dieser Formel jemand einen anderen nur dann als bloßes Mittel an, wenn er ihn einerseits zugunsten seines Eigeninteresses benutzt und wenn er andererseits bereit ist, den anderen alle erdenkbaren Nachteile erleiden zu lassen, die dem Eigeninteresse dienlich sind. Wer dagegen jemanden als Mittel zu moralischen Zwecken gebraucht oder kontrafaktisch bereit wäre, sein eigenes Wohlergehen für den anderen zu opfern, wenn dies die Gesamtlage verbesserte, der verstoße nicht gegen Kants Instrumentalisierungsverbot.

Zur Illustration eignet sich folgendes Gedankenexperiment, das Parfit in Anlehnung an die sogenannten *Trolley-Probleme* entwirft, die Philippa Foot in den 1960er Jahren in die philosophischen Debatten eingebracht hat: Ein führerloser Zug rast unkontrolliert

48 GMS, AA 04, S. 429.

49 Allerdings ist Parfit skeptisch, dass sich daraus eine sinnvolle Bewertung von Handlungen ergeben kann; vgl. OWM, Bd. I, S. 232.

einen Berg hinunter und wird eine Gruppe von fünf Personen überrollen. Die einzige Möglichkeit, dies zu verhindern, bestünde darin, einen Unbeteiligten gegen dessen Willen von einer Brücke auf die Schienen zu stoßen, um damit die automatisierte Notbremse auszulösen. Auch wenn es auf den ersten Blick so aussieht, als würde der Unbeteiligte als bloßes Mittel zur Rettung anderer gebraucht, verstößt eine solche Handlung laut Parfit nicht gegen das Instrumentalisierungsverbot: Der Unbeteiligte würde als Mittel zu einem moralischen Zweck gebraucht, und es kann vorausgesetzt werden, dass der Akteur ihn nur notgedrungen opfert und kontrafaktisch bereit wäre, eigene Opfer in Kauf zu nehmen, um den Unbeteiligten zu entlasten.<sup>50</sup>

Nachdem auf die gleiche Weise weitere Prinzipien und typische Annahmen von Kantianern diskutiert wurden, wendet sich Parfit in Teil III dem ersten Kandidaten für ein oberstes Moralprinzip zu: der Universalisierungsformel des Kategorischen Imperativs. Im Einklang mit einer bedeutenden Richtung innerhalb der Kantforschung<sup>51</sup> hält Parfit zunächst fest, dass Kant zu Unrecht glaubt, die Universalisierungsformel könne zeigen, dass die allgemeine Ausübung strikt verbotener Maximen gar nicht widerspruchsfrei *gedacht* werden kann.<sup>52</sup> Mehr Potential habe demgegenüber der Test, ob ich die allgemeine Ausübung einer bestimmten Maxime bzw. die allgemein geteilte Überzeugung, dass das Handeln nach einer bestimmten Maxime erlaubt sei, *wollen* kann.

Diese Form, die Universalisierungsformel als moralischen Test anzuwenden, könne durchaus zu einem plausiblen obersten Moralprinzip weiterentwickelt werden. Allerdings sei sie einer ganzen Reihe von Einwänden ausgesetzt, die jeweils Modifikationen erforderlich machten. Verkürzt kann Parfits Argumentation wie folgt verstanden werden:<sup>53</sup>

Kant bezieht sich mit der Universalisierungsformel auf Maximen, also subjektive Handlungsgrundsätze von Akteuren. Nun

50 Parfit dehnt diese Diskussion in OWM, Bd. 3, Kap. 56 mit analogen Ergebnissen auf weitere Fassungen des Instrumentalisierungsverbots aus.

51 V. a. im Anschluss an den einflussreichen Aufsatz Herman 1989.

52 Vgl. OWM, Bd. 1, S. 275-284.

53 Parfit diskutiert in OWM, Bd. 1, Kap. 13 und 14 eine Reihe von weiteren Einwänden, die in diesem Band nicht abgedruckt sind, und liefert in Kap. 14 außerdem einen Vergleich der Kantischen Theorie mit der Goldenen Regel.

kann ein Akteur, indem er *einer* Maxime folgt, oftmals sowohl erlaubte als auch verbotene Handlungen durchführen. Folgt er etwa der egoistischen Maxime, immer nur zugunsten seines eigenen Vorteils zu handeln, kann er einerseits seine Steuern bezahlen, weil Steuerhinterziehung eine Strafe nach sich ziehen würde; andererseits aber einem hilfsbedürftigen Menschen die Unterstützung verweigern, weil dies für ihn mit Kosten verbunden wäre. Während es offenkundig moralisch erlaubt ist, seine Steuern zu bezahlen, ist es (zumindest unter bestimmten Umständen) verboten, Hilfestellung zu verweigern. Die egoistische Maxime führt daher einerseits zu erlaubten Handlungen (Steuern zu zahlen), andererseits zu verbotenen (Hilfe zu verweigern). Analog dazu scheint in den meisten Fällen zu gelten, dass die verschiedenen Handlungen, die auf Basis *einer* Maxime ausgeführt werden, teils erlaubt und teils verboten sind. Da die Ethik danach frage, wie wir handeln sollen, helfe es uns nicht weiter, wenn wir *Maximen* auf ihre moralische Zulässigkeit hin testen: Dieser Test habe keine relevante Aussagekraft. Parfit nennt dies den *Einwand uneinheitlicher Maximen*.<sup>54</sup> Statt Maximen zu bewerten, sei danach zu fragen, welche *Arten von Handlungen* allgemein gewollt werden können.

Doch wie müssen diese Arten von Handlungen definiert werden, damit sich das Problem nicht wiederholt? Offenbar müssen sie hinreichend genau beschrieben werden, um erlaubte von verbotenen Handlungsarten unterscheiden zu können. Doch würde man Handlungsarten zu detailliert beschreiben, fiel die allgemeine Ausübung der Handlung mehr oder weniger mit ihrer einmaligen Ausführung zusammen. Nehmen wir beispielsweise an, jemand sieht eine Frau in einem weißen Kleid in einem Café sitzen, die gerade Erdbeeren isst und die letzte Seite von Spinozas *Ethik* liest. Nun könnte er sich theoretisch fragen, ob die allgemeine Ausübung der Handlung »einer Frau in einem weißen Kleid, die gerade Erdbeeren isst und die letzte Seite von Spinozas *Ethik* liest, den Geldbeutel stehlen« vernünftigerweise gewollt werden kann. Er müsste dies klarerweise bejahen: Ihm würde die allgemeine Ausübung einen großen Vorteil verschaffen, und es gäbe nur einen schwachen unparteiischen Grund, der dagegen spräche, weil ein solcher Fall wohl kein zweites Mal aufträte.

54 Vgl. in diesem Band, S. 282 f. bzw. OWM, Bd. 1, S. 292 f.

Als Konsequenz dieses *Seltenheitseinwands*<sup>55</sup> fordert Parfit, dass alle und nur die moralisch relevanten Eigenschaften zur Beschreibung einer Handlung herangezogen werden müssen; im genannten Beispiel würde die Farbe des Kleides und die Beschreibung einer nebensächlichen Tätigkeit offenkundig nicht dazugehören. Auf der anderen Seite garantiert die Einbeziehung *aller* moralisch relevanten Eigenschaften, dass die fragliche Art von Handlung *immer* erlaubt oder *immer* verboten ist, es also keine uneinheitliche Bewertung gibt, wie sie im Fall der Maximenprüfung vorlag.

Doch auch bei Einhaltung dieser Bedingung sind viele Handlungstypen so beschaffen, dass der Akteur nie befürchten muss, dass er selbst einmal negative Auswirkungen ihrer allgemeinen Ausübung zu spüren bekäme. Dies trifft insbesondere auf Handlungstypen zu, die eine diskriminierende Wirkung haben: Ein wohlhabender Mensch könnte durchaus vernünftigerweise wollen, dass sozial schwache Menschen wirtschaftlich ausgenutzt werden, weil er selbst davon profitiert. Um diesem *Einwand der Unumkehrbarkeit*<sup>56</sup> zu begegnen, muss die Universalisierungsformel derart reformuliert werden, dass sie nicht mehr danach fragt, ob *ich* die allgemeine Ausübung eines Handlungstyps wollen kann, sondern ob *jedermann* seine allgemeine Ausübung wollen kann.

Wenn Handlungstypen auf diese Weise geprüft werden können, lassen sich Moralprinzipien mittlerer Ebene formulieren, die diese Handlungstypen erlauben oder verbieten. Anstatt über Handlungstypen zu reden, die mit allen moralisch relevanten Eigenschaften beschrieben werden, kann man daher kürzer über die Prinzipien reden, die solche Handlungen bewerten. Damit ist Parfit bei einer Formel angelangt, die gleichzeitig auch die grundlegenden Erfordernisse des Kontraktualismus erfüllt und daher als *Formel des Kantischen Kontraktualismus* bezeichnet werden kann:

Jeder sollte denjenigen Prinzipien folgen, deren allgemeine Akzeptanz jeder vernünftigerweise wollen kann.<sup>57</sup>

Die überraschendste Pointe der *Triple Theory* besteht nun in Parfits Nachweis, dass sich anhand dieses obersten Moralprinzips exakt

55 Vgl. in diesem Band, S. 277 f. bzw. OWM, Bd. I, S. 289.

56 Vgl. in diesem Band, S. 293 f. bzw. OWM, Bd. I, S. 334 ff.

57 In diesem Band, S. 304; vgl. OWM, Bd. I, S. 342.

die gleichen untergeordneten Moralprinzipien ergeben, die auch im Rahmen einer plausiblen Variante des Konsequentialismus vertreten werden. Dem Konsequentialismus zufolge hängt die Frage, ob unsere Handlungen gut oder schlecht sind, allein von Tatsachen darüber ab, wie die Welt den besten Verlauf nehmen würde. Handlungskonsequentialisten wenden dieses Prinzip direkt auf einzelne Handlungen an: Jede Handlung wird daraufhin geprüft, ob sie im Vergleich zu alternativen Handlungen den besten Weltverlauf herbeiführt. Regelkonsequentialisten wenden den konsequentialistischen Test dagegen auf Regeln an: Geprüft wird, welche Regel oder welches Set an Regeln, würden sie von allen befolgt, den besten Weltverlauf herbeiführten. Parfit bezeichnet eine Regel, die den besten Weltverlauf herbeiführt, wenn jeder ihr folgt, in Anlehnung an das Begriffsinstrumentarium von William David Ross<sup>58</sup> als »*optimific*«, was im Deutschen am ehesten mit »optimal« wiedergegeben werden kann.

Der *Handlungskonsequentialismus* würde viele Handlungen zulassen, die unseren Intuitionen nach eindeutig falsch sind. So würde etwa ein Arzt die Pflicht haben, einen Patienten gegen dessen Willen heimlich zu töten, wenn er mit dessen Organen fünf anderen Patienten das Leben retten könnte. Für den *Regelkonsequentialismus* gilt das nicht, denn dieser berechnet auch die negativen »Nebenwirkungen« einer allgemeinen Regelbefolgung mit ein: Würde eine Regel allgemein anerkannt, der zufolge jeder Arzt einen Patienten zu Gunsten einer größeren Zahl anderer Patienten töten würde, sofern er dies unbemerkt vollbringen kann, so würde sich kaum noch jemand einem Arzt anvertrauen. Als Folge wären wir alle medizinisch schlechter versorgt.<sup>59</sup>

58 Vgl. Ross 2002 [1930], S. 34f.

59 Vgl. OWM, Bd. 1, S. 363. In RaP war von einer Ablehnung des Handlungskonsequentialismus noch nichts zu sehen. Die reduktionistische Theorie personaler Identität legt sogar nahe, dass jeder von uns gute Gründe hätte, der beschriebenen tödlichen Organspende zuzustimmen, denn auf den Fortbestand des *eigenen* Lebens kommt es ihr zufolge nicht an. Konsequenterweise *verteidigt* Parfit den Handlungskonsequentialismus in RaP in einigen Hinsichten, insbesondere gegen den Einwand, dass es für die Ziele des Handlungskonsequentialismus nicht gut wäre, wenn alle Menschen zu Handlungskonsequentialisten würden, weshalb der Handlungskonsequentialismus »selbsterstörerisch« [*self-defeating*] sei. Zuletzt versucht Parfit in OWM, Bd. 3, den Handlungskonsequentialismus ein Stück weit mit unseren moralischen Intuitionen zu versöhnen (Kap. 55 und 57),

Daher konzentriert sich Parfit auf den Regelkonsequentialismus und die damit verbundenen »optimalen« Prinzipien. Er möchte zeigen, dass die *Formel des Kantischen Kontraktualismus* im Ergebnis ebenfalls nur optimale Prinzipien für gültig befindet.<sup>60</sup> Optimale Prinzipien seien also genau diejenigen Prinzipien, deren allgemeiner Ausübung jedermann vernünftigerweise zustimmen kann. Die beiden wesentlichen Prämissen des Arguments lauten, dass jedermann starke *unparteiische* Gründe hat, optimalen Prinzipien zuzustimmen, und dass es keine signifikanten parteiischen Gründe gibt, die nicht-optimale Prinzipien für *einen* Personenkreis als zustimmungsfähig auszeichnen, ohne dass sie dadurch ihre Zustimmungsfähigkeit in den Augen *anderer* Personen verlieren.

Der wichtigste Prüfstein dieser Argumentation sind kategorische Verbotsregeln, die in der kantischen Tradition meist vertreten, von Konsequentialisten aber verworfen werden. Nehmen wir eine Regel, die es im oben vorgestellten *Trolley-Problem* erlauben würde, jemanden zur Rettung von fünf anderen Personen vor den Zug zu stoßen. Eine solche Regel wäre optimal, steht aber der kantischen Intuition des kategorisch verbotenen Tötens von Unschuldigen entgegen. Ein unparteiischer Beobachter könnte dieser Regel vernünftigerweise zustimmen, denn sie minimiert den Schaden, den die Welt durch das Vorkommnis erfährt. Doch könnte auch die Person, die auf die Schienen gestoßen wird, der Regel zustimmen? Sie hätte im Prinzip starke parteiische (selbstbezogene) Gründe, diese Regel abzulehnen. Doch für jede Regel, die eine solche Handlung untersagt, würde gelten, dass die fünf anderen Personen starke parteiische Gründe haben, sie abzulehnen. Wenn also überhaupt eine Regel als allgemein zustimmungsfähig ausgewiesen werden kann, dann die optimale, der ja immerhin jeder Unbeteiligte zustimmen kann.

Könnte es vorkommen, dass gar keine Regel *jedermanns* Zustimmung erhalten kann? Um diesen Fall auszuschließen, bleibt nachzuweisen, dass derjenige, der auf die Schienen gestoßen wird, der optimalen Regel vernünftigerweise zustimmen *kann* (auch wenn eine Person in dieser Lage faktisch vermutlich nicht zustimmen *möchte*). Hierbei wird die Idee der Verallgemeinerung beson-

bringt dann aber letztlich starke Argumente zugunsten des Regelkonsequentialismus vor (Kap. 58).

60 Vgl. OWM, Bd. I, S. 378.

ders wichtig, die dem Kategorischen Imperativ zugrunde liegt und die laut Parfit wie ein Vergrößerungsglas wirkt.<sup>61</sup> Die allgemeine Ausübung der fraglichen Regel würde nicht nur in diesem Fall, sondern auch in vielen künftigen Fällen die Zahl der Opfer minimieren. Die unparteiischen Gründe, der Regel zuzustimmen, sind angesichts dieser Verallgemeinerung extrem stark, während die parteiischen Gründe des konkreten Opfers unverändert bleiben. Im Ergebnis erscheint es somit nicht unplausibel, dass die unparteiischen Gründe die parteiischen Gründe aufwiegen (wenn auch vielleicht nicht überwiegen) können. Eine Zustimmung des Opfers wäre also möglich, ohne gegen die Bedingung der Vernünftigkeit zu verstoßen.

Obwohl Parfit andeutet, dass er das fragliche Problem tendenziell so lösen würde,<sup>62</sup> möchte er seine Theorie nicht auf solche substantiellen ethischen Aussagen festlegen und skizziert, wie vielleicht auch innerhalb seiner Theorie kategorische Verbote wie das Töten Unschuldiger begrifflich integriert werden könnten: Es ist denkbar, dass die Tötung eines Unschuldigen intrinsisch so schlecht ist, dass die Welt nicht mehr auf bestmögliche Weise verlief, wenn diese Handlung ausgeübt würde.<sup>63</sup> Ob Parfit damit unseren Intuitionen ausreichend gerecht werden kann, bleibt selbstredend fraglich (siehe Abschnitt 2.5).

Das schwierige Unterfangen, die inhaltliche Übereinstimmung von Kantianismus und Konsequentialismus nachzuweisen, ist in den Augen Parfits damit jedoch geglückt. Um das »Triple« zu vervollständigen, ist Scanlons Variante des Kontraktualismus noch hinzuzufügen. Ohne diesen Schritt an dieser Stelle weiter erläutern zu können, ergibt sich als vollständige Formulierung der *Triple Theory*:

Eine Handlung ist genau dann falsch, wenn solche Handlungen von einem Prinzip verboten werden, das

(1) zu denjenigen Prinzipien gehört, die den besten Weltverlauf herbeiführen würden, wenn sie allgemeine Gesetze wären,

61 Vgl. OWM, Bd. 1, S. 307.

62 Vgl. OWM, Bd. 2, S. 154.

63 Vgl. in diesem Band, S. 335-338 bzw. OWM, Bd. 1, S. 394f.

(2) zu den einzigen Prinzipien gehört, von denen jeder wollen kann, dass sie allgemeine Gesetze werden,

und das

(3) niemand vernünftigerweise zurückweisen kann.<sup>64</sup>

## 2.5 Die Rezeption der Moralphilosophie

Mit seinen Überlegungen zum Problem der Nicht-Identität, zur Abstoßenden Schlussfolgerung und zum Prioritarismus hat Parfit große Debatten angestoßen, die ihn innerhalb der praktischen Philosophie zu einem der meistdiskutierten Autoren der Gegenwart machen – wenngleich alle drei Debatten bislang selten in deutscher Sprache rezipiert werden. Zudem hat sich die Diskussion in einigen Bereichen verselbstständigt, so dass nicht mehr durchgehend ein ausdrücklicher Bezug zu Parfit hergestellt wird.

Parfits Egalitarismuskritik hatte großen Einfluss,<sup>65</sup> und der Prioritarismus ist laut Brad Hooker mittlerweile sogar »die dominierende Form des Konsequentialismus«<sup>66</sup> geworden. Davon unbeschadet sind die Kontroversen um Gleichheit und Vorrangigkeit noch längst nicht abgeschlossen.<sup>67</sup> Das Problem der Nicht-Identität findet in den Debatten um die Klimaethik<sup>68</sup> und historische Ungerechtigkeiten<sup>69</sup> bis heute Beachtung. Die Populationsethik hat sich mittler-

64 OWM, Bd. I, S. 412 f. (Übers.: Hg.).

65 Zu Teilen dokumentiert in Krebs 2000.

66 Hooker 2011 (Übers.: Hg.).

67 Zur aktuellen Debatte vgl. insbesondere die Beiträge in *Economics and Philosophy* 31/2 (2015); für explizite Auseinandersetzungen mit Parfit vgl. Persson 2012; Otsuka 2012 und Dufner (i. E.).

68 Vgl. z. B. Broome 1992; Gesang 2011, S. 136 ff.; Roser/Seidel 2015, S. 29. In der Klimaethik werden darüber hinaus Parfits Beiträge zu den Fragen behandelt, ob zeitlich entferntere Auswirkungen unserer Handlungen weniger Gewicht bekommen sollten als die zeitlich näher liegenden (vgl. Parfit 1983 und etwa Attfield 2014) und warum einzelne klimaschädliche Handlungen moralisch verwerflich sind, obwohl sie, je für sich betrachtet, keinen messbaren Einfluss auf das Klima haben können (vgl. RaP, § 27).

69 Um nur einige Beispiele herauszugreifen: Cohen 2009, Roberts/Wasserman 2009 und für den deutschsprachigen Raum Meyer 2006. Für eine Darstellung der Debatte vgl. Roberts 2015.

weile als eigene Subdisziplin der Angewandten Ethik etabliert und geht, über Parfits Problembeschreibung hinausreichend, etwa der Frage nach, welche Pflichten wir bei drohender Über- oder auch Unterbevölkerung der Welt haben könnten. Kann es beispielsweise eine moralische Pflicht geben, Kinder zu zeugen?<sup>70</sup> Die Debatte um die Abstoßende Schlussfolgerung wird dagegen nach wie vor in relativ enger Anlehnung an Parfits Texte geführt, insbesondere mit Bezug auf die wirkmächtige Problembeschreibung in RaP.<sup>71</sup> Aber auch Parfits jüngster Vorschlag, die Abstoßende Schlussfolgerung zu vermeiden, ist bereits intensiv diskutiert worden.<sup>72</sup>

Im Gegensatz zu den moralphilosophischen Themen des Frühwerks hat die *Triple Theory* bislang keinen vergleichbaren Erfolg hervorgerufen. Bis heute ist jedenfalls kein grundlegender Paradigmenwechsel im Streit zwischen Kantianern und Konsequentialisten wahrzunehmen; die Landkarte der philosophischen Streitpunkte ist nicht neu zu zeichnen, wie es sich Parfit noch in den *Tanner Lectures*<sup>73</sup> erhofft hatte. Zahlreiche Reaktionen auf die *Triple Theory* – vier Kommentare, die sich auf die *Tanner Lectures* beziehen, sind bereits im zweiten Band von OWM abgedruckt – sind eher kritisch: Parfits Methode, einzelne Prinzipien anhand des Common Sense bzw. anhand unserer Intuitionen zu Gedankenexperimenten zu prüfen, anstatt sie aus übergeordneten Überlegungen zu deduzieren, wird ebenso kritisiert wie seine angeblich zu große Bereitschaft, deontologische Überzeugungen zugunsten der Vereinbarkeit mit alternativen Theoriemodellen aufzugeben.<sup>74</sup>

Ein Standardeinwand gegen Parfits These, dass wir regelkonsequentialistische Prinzipien akzeptieren sollten, beruft sich darauf, dass es einzelnen Akteuren unmöglich ist, auch nur annähernd herauszufinden, welche Prinzipien letztlich den besten Weltverlauf herbeiführen würden; daher sei es Akteuren letztlich unmöglich, aus Einsicht moralisch richtig zu handeln.<sup>75</sup> Parfit wehrt sich gegen diesen Vorwurf, indem er darauf verweist, dass die *Triple Theory* lediglich beweisen soll, dass es keinen Widerspruch zwischen den

70 Vgl. Blackorby/Bossert 2005; Goldin 2014; Roberts 2011.

71 Arrhenius u. a. 2014.

72 Arrhenius 2016; Chang 2016.

73 Parfit 2004b, S. 368.

74 Wood 2011; Herman 2011; Scruton 2014; Fried 2012; Scheffler 2011, S. xxvii.

75 Scruton 2014, S. 633.

gängigen Moraltheorien gibt; sie diene aber nicht dazu, Akteuren in konkreten Situationen einen Maßstab zu bieten.<sup>76</sup> Die Fragen, auf die es in den konkreten Entscheidungen von realen Akteuren ankomme, seien nämlich ohnehin nicht strittig, wenn einmal sichergestellt ist, dass es überhaupt auf etwas ankomme: Dass wir die Erde kommenden Generationen als bewohnbaren Ort überlassen oder Armut bekämpfen sollen, erfordert keine aufwändige Bestimmung der optimalen Prinzipien.

Die größte Hürde für Kantianer, regelkonsequentialistische Prinzipien zu akzeptieren, scheint aber in der Unmöglichkeit zu liegen, kategorische Verbote, wie sie etwa aus einigen Interpretationen der Menschenwürde abzuleiten sind, regelkonsequentialistisch zu rekonstruieren. Der moralische Status von Personen sowie deren Autonomie blieben demzufolge in der *Triple Theory* unzureichend berücksichtigt, weil letztlich all das, was als Verstoß gegen die Idee von Personen als autonomen Akteuren zu beschreiben sei, durch die Generierung ausreichend guter Folgen legitimiert werden könnte.<sup>77</sup>

Wenngleich Parfit wirkmächtig Problemstellungen in die moralphilosophische Debatte eingeführt hat, die uns noch lange beschäftigen werden, bleiben seine Lösungsvorschläge also strittig. Was OWM betrifft, ist es allerdings noch zu früh, um eine Prognose über die mittelfristige Resonanz eines Werkes abgeben zu können, das bereits kurz nach Erscheinen sogar von Parfits Kritikern als »Meilenstein der Philosophie«<sup>78</sup> und – in Anspielung auf Parfits Wertschätzung für Sidgwick – als das wichtigste Buch der Ethik seit 1873, also seit Erscheinen von dessen *Methods of Ethics*, bezeichnet worden ist.<sup>79</sup>

### 3. Metaethik und Theorie der Normativität

Während sich Parfit bereits seit den Zeiten von RaP auch mit metaethischen Problemen auseinandergesetzt hat, kulminieren seine jahrzehntelangen Bemühungen in dem 2011 erschienenen OWM,

76 Parfit 2013, S. 31f.

77 Scruton 2014; Hoesch/Sticker (i. E.); Otsuka 2009; Herman 2011; Wood 2011; Ross 2009; Darwall 2014, S. 99-103.

78 Scruton 2014, S. 621.

79 Singer 2011.

auf das wir bei der Dokumentation seiner Theorie der Normativität in diesem Band unser Hauptaugenmerk legen. Parfit gliedert seine Metaethik dort in eine Theorie der Gründe (1. Teil) und eine Theorie ethischer Wahrheit (6. Teil) und ergänzt das im dritten Band in mancherlei Hinsicht. Dabei verteidigt er einen Ansatz, der Gründe als objektive, wertbasierte Tatsachen versteht. Diese Gründe werden als irreduzibel normativ angesehen: Weder können wir mit nicht-normativen Begriffen über sie Aussagen treffen, noch sind Gründe weiter mittels Tatsachen analysierbar, die selbst nicht normativ sind.

Wenn Parfit von »nicht-normativen« Begriffen und Eigenschaften spricht, so bezieht er sich auf solche, die von den Natur- und Sozialwissenschaften verwendet bzw. beschrieben und als »wertfrei« betrachtet werden. Philosophische Positionen, die sich in ihrer Erklärung von Phänomenen weitgehend auf solche nicht-normativen Eigenschaften beschränken, werden oft mit dem Attribut »naturalistisch« versehen; im Gegenzug nennt man Positionen, die irreduzibel normative Begriffe verwenden, oft »nicht-naturalistisch«.

Auf den ersten Blick scheint diese Beschreibung Parfit kurz und bündig als *nicht-naturalistischen Realisten* auszuzeichnen: Die Existenz irreduzibel normativer Begriffe und Eigenschaften sollte Letztere, so der Leitgedanke des Realismus, unproblematisch als wirklich oder *real* verbürgen. Jedoch bezeichnet Parfit seine Position überraschenderweise als *nicht-realistischen Kognitivismus* [*Non-realist Cognitivism*] und grenzt sich damit explizit von dieser Interpretation ab: Zwar beziehen wir uns mit normativen Begriffen auf normative Wahrheiten (insofern ist Parfits Position »kognitivistisch«), diesen kommt aber keine gleich näher zu erläuternde »ontologische Wirklichkeit« zu (das macht Parfits Position »nicht-realistisch«) Dennoch erweisen sich die Unterschiede von Parfits Position, wie wir gleich sehen werden, gegenüber konkurrierenden *nicht-realistischen* Positionen als viel bedeutender als gegenüber seinen realistischen Mitbewerbern.

Parfits Abgrenzung innerhalb der verschiedenen Varianten des metaethischen Realismus ist gegenüber seiner Abgrenzung zu den nicht-realistischen Positionen differenziert zu betrachten. Dessen Kernaussage – es gibt ethische Tatsachen, über die wir mit unseren ethischen Sätzen Aussagen treffen – kann nämlich auch von einer naturalistischen Position bejaht werden. Tatsächlich grenzt sich der

naturalistische Realismus von seinen nicht-naturalistischen Fassungen gerade durch die Ablehnung der Behauptung ab, Normativität sei *ontologisch wie semantisch irreduzibel oder fundamental*.

Daraus folgt aber: Für den nicht-naturalistischen Realismus, der die Fundamentalitätsthese mitträgt, ergibt sich aus der obigen Beschreibung alleine noch kein wirkliches Abgrenzungspotential zum nicht-realistischen Kognitivismus. Warum dann Parfits starke Abgrenzung? Parfit sieht seine eigene Position trotz der Rede von normativen Tatsachen von einer entscheidenden Verpflichtung des Realismus noch unberührt: diejenige einer *ethischen Wirklichkeit*. Demnach behaupten Realisten nicht nur, dass unsere ethischen Aussagen unabhängig von unseren Einstellungen des Wünschens und Meinens wahr sind, sondern auch, dass ihre Wahrheit von einer ethischen Wirklichkeit (ontologisch) getragen und (epistemisch) bestätigt wird.

Eine solche Anbindung an die Wirklichkeit lässt sich für Parfit aus dem bloßen Begriff der Tatsache nicht gewinnen. Er unterscheidet folgerichtig zwischen Tatsachen, die ontologisch »gewichtig« sind, und solchen, die das nicht sind.<sup>80</sup> Als Beispiele für Tatsachen ohne ein derartiges *fundamentum in re* nennt Parfit mathematische und logische Wahrheiten: Dass zwei und drei fünf ergibt, verpflichtet nicht auf die Annahme der Existenz von Zahlen. Zur Verdeutlichung fordert Parfit den Leser auf, er möge sich vorstellen, das Universum hätte nie existiert. Viele Wahrheiten über wirkliche Sachverhalte wären dann nicht länger gültig, etwa die Tatsache, dass die Erde um die Sonne kreist. Die gerade erwähnte Gleichung würde aber, zusammen mit den anderen Tatsachen der Mathematik, ihre Gültigkeit behalten. Auch wenn nichts existiert, können folglich Tatsachen bestehen, und nicht jede Wahrheit ist von etwas abhängig, das wirklich besteht.<sup>81</sup>

### 3.1 Parfits Gründe-Objektivismus

Welche Gründe aber haben wir, normative Tatsachen dem ontologisch »ungewichtigen« Bereich zuzuordnen, die selbst dann be-

<sup>80</sup> Wobei er darüber hinaus Wahrheiten über mögliche Sachverhalte zulässt (beispielsweise die Tatsache, dass es Einhörner geben könnte), die einen ontologisch *weniger gewichtigen* Status haben. Vgl. Parfit OWM, Bd. 2, S. 481 und S. 727 f.

<sup>81</sup> Vgl. OWM, Bd. 2, S. 482.

stunden, wenn es die ontologisch gewichtigen Dinge, von denen sie handeln, nicht gibt? Nehmen wir etwa an, das Universum hätte es nie gegeben. Ist es offenkundig, dass die Norm, dass man Wesen mit menschlichen Eigenschaften nicht zum Spaß umbringen darf, dennoch unverändert besteht?

Parfitts Antwort darauf markiert einen *zweiten* wichtigen Unterschied zu vielen nicht-naturalistischen Realisten. Zwar bezeichnet er seine eigene Theorie normativer Gründe, wie zu Beginn erwähnt, als »wertbasiert«. Zugleich macht er jedoch klar, dass er dabei die Existenz von Tatsachen über Gründe nicht auf die Existenz von Tatsachen über Werte stützen möchte – Tatsachen darüber, welche Dinge gut oder schlecht sind –, die neben der Welt der Gründe eine zweite fundamentale Kategorie des Normativen bilden würden. Vielmehr sieht er, ähnlich wie Thomas Scanlon und Thomas Nagel,<sup>82</sup> Wert-Tatsachen selbst durch Tatsachen über Gründe gerechtfertigt:

Den wertbasierten Theorien zufolge haben wir sehr wohl [...] Gründe. Diese Gründe werden von verschiedenen natürlichen Merkmalen der Gegenstände unserer Wünsche geliefert, und von diesen Gründen erhalten alle anderen Gründe ihre Kraft. Diese Aussage [...] macht keinen Gebrauch von den Begriffen des *Guten* oder *Wertvollen*. Wenn ein Gegenstand solche gründegebenden Eigenschaften aufweist, können wir ihn als »gut« bezeichnen, aber das ist bloß eine Abkürzung.<sup>83</sup>

In den ersten drei Kapiteln von OWM, die im vorliegenden Band unter dem Titel »Gründe« abgedruckt sind, stellt Parfit seine Form des *Gründe-Objektivismus* verschiedenen *subjektivistischen Theorien über Gründe* gegenüber. Kennzeichnendes Merkmal subjektivistischer Theorien ist die Abhängigkeit der Gültigkeit ihrer Gründe von persönlichen Einstellungen des Wollens oder Wünschens. Je nach subjektivistischer Theorie werden diese dabei unterschiedlich stark »idealisiert« und müssen verschiedenen Anforderungen genügen. Im Ergebnis kommt es jedoch bei der Beantwortung der

82 Vgl. Scanlon 1998 und Nagel 2013, Kap. 5.

83 Parfit 2001, S. 20. (Übers.: Hg.).

Frage, für welches Verhalten wir Gründe haben, stets auf Tatsachen über diese *subjektiven Einstellungen* an, nicht auf diejenigen über *ihre Gegenstände*. Ob ich Gründe habe, ein Medikament gegen meine Kopfschmerzen zu nehmen, ist also beispielsweise nicht von der Tatsache abhängig, dass dessen schmerzlindernde Wirkung dafür spricht, es zu nehmen, sondern davon, ob ich, sofern ich wohlinformiert bin und weiß, welche Wirkungen dessen Einnahme auf meine Schmerzen hat, den Wunsch ausbilden würde, es einzunehmen.

Im Rahmen seines Objektivismus argumentiert Parfit für die entgegengesetzte Position: Gründe-Tatsachen beziehen sich auf die Gegenstände unserer Einstellungen – genauer: auf die über Tatsachen erfassbaren, nicht-normativen Eigenschaften bestimmter *Ereignisse* oder *Zustände* –, freilich ohne mit diesen identisch zu sein. Dass die nicht-normativen Eigenschaften bestimmter Tatsachen für oder gegen ein bestimmtes Verhalten sprechen, wird daher als *zusätzliche*, normative Tatsache über solche Ereignisse und Zustände verstanden. Im Unterschied zur Analyse der subjektivistischen Theorien sorgt sie erst dafür, *dass wir Gründe haben* und dadurch gerechtfertigt sind, sie zum Gegenstand unseres Wollens zu machen, um uns dementsprechend zu verhalten. In Abgrenzung dazu beruht die *Rationalität* unseres Verhaltens auf unseren *normativen Überzeugungen* darüber, was Gegenstand unseres Wünschens und Wollens sein sollte. Dadurch kann es vorkommen, dass wir uns rational verhalten, ohne dasjenige zu tun, wofür wir *wirklich* Gründe haben.<sup>84</sup>

Parfit sieht die Vorteile des Objektivismus gegenüber dem Subjektivismus *erstens* dadurch begründet, dass selbst in Fällen, in denen persönliche Einstellungen des Wollens und Wünschens uns *unmittelbar* Gründe zu geben scheinen, diese tatsächlich bestenfalls *mittelbar* aktiv werden. *Zweitens* führt er Problemfälle auf, bei denen subjektivistische (im Gegensatz zu objektivistischen) Theorien zu kontraintuitiven Lösungen führen.

Der *erstgenannte* Vorteil wird von Parfit an hedonischen Gründen veranschaulicht, also Gründen, die sich direkt auf ein Lust- und Unlustempfinden beziehen: Unsere Vorlieben und Abneigungen sind einerseits Bewusstseinszustände; gleichzeitig scheinen wir

84 Vgl. neben OWM, Bd. I, Kap. 5 auch Parfit 1997b, S. 99 und Parfit 2001, S. 25.

häufig unser Verhalten zu begründen, indem wir uns auf derlei Zustände beziehen, etwa wenn wir den Kauf unseres Lieblingseises rechtfertigen oder jemandem unsere ablehnende Reaktion auf einen langweiligen Film erklären. Parfit argumentiert nun, dass auch in diesen Fällen *nicht* die Tatsachen, dass wir bestimmte von uns gemochte oder abgelehnte Bewusstseinszustände realisieren oder verhindern *wollen* oder uns dies *wünschen*, unsere Gründe dafür erst »machen«. Unser Wollen und Wünschen wird vielmehr selbst über normative Tatsachen gerechtfertigt, nämlich diejenigen, dass wir Gründe haben, Zustände herbeizuführen, die gemocht werden, und solche zu verhindern, denen wir ablehnend gegenüberstehen. Was uns einen normativen Grund gibt, ist also etwa der erwartete *Genuss*, den das Eis hervorruft, und nicht der *Wunsch* danach.

Dieser Argumentationslinie folgend, ergibt sich der *zweite* von Parfit genannte Vorteil eines Objektivismus: Die Tatsache, dass bestimmte Dinge für ein Verhalten sprechen, erweist sich gegenüber der Ausbildung von Wünschen oder willentlichen Zielen als stabil. In einem von Parfit häufig benutzten Beispiel bildet eine Person die Präferenz aus, zukünftige Schmerzen zwar allgemein abzulehnen, ihnen jedoch indifferent gegenüberzustehen, sofern sie an einem künftigen Dienstag auftreten. Diese Präferenz nennt Parfit »irrational« (und kann sich hier vermutlich unserer Unterstützung sicher sein), denn Schmerz sei Schmerz: Alle zukünftigen Schmerzen sollten gemäß Sidgwick's »Axiom der temporalen Unparteilichkeit« hinsichtlich ihrer zeitlichen Entfernung gleich wichtig für uns sein.<sup>85</sup> Ein Subjektivist könne aber nicht erklären, weshalb diese Präferenz irrational sein könnte: Wenn alle Gründe von subjektiven Einstellungen abhängen, können diese Einstellungen selbst nicht wieder anhand von Gründen überprüft werden.

### 3.2 Die Fundamentalität des Normativen

In Parfits Vergleich seines nicht-realistischen Kognitivismus mit anderen Metaethiken lässt sich eine deutliche Zweiteilung zwischen Positionen ausmachen, die – wie er selbst – die Fundamentalität des Normativen zugestehen, und solchen, die das nicht tun. Dabei setzt Parfit in seiner Auseinandersetzung einen deutlichen Schwer-

<sup>85</sup> Vgl. dazu auch S. 232f. von »Normative Wahrheiten«, dem hier abgedruckten Kap. 30 aus OWM, Bd. 2.

punkt auf die zweitgenannte Gruppe von Positionen, die er als dem normativen Nihilismus nahestehend charakterisiert: Würden sie richtigliegen, wäre nichts von Bedeutung – *nothing would matter*. Diese Gruppe wird von Parfit weiter unterteilt in naturalistische Realisten verschiedener Couleur einerseits und Non-Kognitivisten andererseits, wobei er innerhalb der letztgenannten Theoriefamilie besonders quasi-realistische Expressivisten diskutiert.

Beginnen wir jedoch zunächst mit einem kurzen Blick auf die Position des nicht-naturalistischen Realisten. Ebenso wie Parfit akzeptieren diese Realisten die Fundamentalitätsthese und damit die begriffliche und ontologische Eigenständigkeit des Normativen. Die Vorteile des nicht-realistischen Kognitivismus ihnen gegenüber ergeben sich im Wesentlichen aus der bereits angesprochenen, von Parfit postulierten ontologischen »Leichtgewichtigkeit« seiner eigenen Position. Demgegenüber blieben Umfang und Art der von den nicht-naturalistischen Realisten eingegangenen ontologischen Verpflichtungen unklar: Während wir wissen, was es heißt, dass es Berge, Katzen und Atome gibt (und man daher einem naturalistischen Realisten keine »merkwürdigen« ontologischen Verpflichtungen vorwerfen kann), geraten wir in Schwierigkeiten, allgemeine Kriterien für die Existenz von normativen Dingen anzugeben. Parfit übernimmt hier eine »No Clear Question View« und meint, dass der nicht-naturalistische Realist selbst nicht wirklich weiß, wovon er spricht.<sup>86</sup>

Diese Strategie weist einige Ähnlichkeiten zum Quasi-Realismus als wichtigster Variante des Non-Kognitivismus sowie zu Spielarten des Quietismus auf,<sup>87</sup> für die die Rede von normativen Tatsachen und deren »objektive« Deutung ebenfalls keinerlei realistisch-ontologische Implikationen aufweisen muss. Im Gegensatz zu beiden Konkurrenten möchte Parfit diese Diskussion jedoch nicht auf der Ebene der normativen Ethik entscheiden. Die Frage, welche metaphysischen Belastungen für die eigene Theorie gelten, lässt sich nicht unter Verweis auf die »robusten« Forderungen, die wir in einigen normativen Diskursen, namentlich dem moralischen, aneinander stellen, einfach abweisen, wie dies Quietisten und Quasi-Realisten zu tun versuchen. Ihnen zufolge ist es *innerhalb* unseres normativen Diskurses eine ausgemachte Sache, dass wir die

86 Vgl. hierzu OWM, Bd. I, S. 476-479.

87 Vgl. etwa Blackburn 1993, Kap. 9 und Dworkin 2011, Kap. 2 f.

Gültigkeit etwa des Folterverbots nicht davon abhängig machen, ob jemand dem persönlich zustimmt oder nicht. Demgegenüber sei es eine *andere* Frage, auf welcher ontologischen Basis wir *außerhalb* dieses Diskurses unsere normativen Überzeugungen bilden würden. Anders Parfit: Für ihn weist die Frage nach der Gültigkeit unserer moralischen Überzeugungen geradewegs in die Metaethik. Innerhalb der Metaethik lassen sich dementsprechend auch abseits unseres alltäglichen Sprechens unterschiedliche normative Phänomene ontologisch unterschiedlich behandeln: Zumindest »dünne«, also deskriptiv unterbestimmte Werte, zum Beispiel das Gute, spielen für Parfit, wie gesehen, keine eigenständige Rolle innerhalb des Normativen, sondern fungieren als sprachliche Abkürzungen für das Vorliegen von Gründen.

Kommen wir damit zu Parfits Diskussion des naturalistischen Realismus und des Expressivismus als der von ihm diskutierten Form des Non-Kognitivismus, die – wie oben vorweggenommen – einen viel größeren Raum einnimmt als die Abgrenzung zu Theorietypen, die die Fundamentalitätsthese mittragen. Abgesehen von der Herausarbeitung ihrer jeweiligen Schwächen, die die ersten sechs Kapitel des 6. Teils von OWM füllt, lässt sich die Verteidigung seiner eigenen Position ihnen gegenüber grob in zwei Richtungen unterteilen: Die erste versucht analog zu der gerade betrachteten Strategie gegenüber dem nicht-naturalistischen Realisten nachzuweisen, dass diejenigen, die den eigenständigen Gehalt des Normativen rundheraus leugnen, häufig nicht wissen, wovon sie sprechen (dazu gleich mehr). Die zweite bemüht sich, darauf aufbauend, ontologische und epistemische Zweifel an der Faktizität des Normativen auszuräumen; hierzu kommen wir in den Abschnitten 3.4 und 3.5.

Die *erste* dieser Strategien wird in dem hier abgedruckten OWM-Kapitel »Normative Wahrheiten« verfolgt, findet sich aber auch bereits in dem 2006 erschienenen Beitrag »Normativity«. Sie wird von Parfit dort wie folgt umrissen:

Viele [...] Autoren ignorieren die Möglichkeit, dass es normative Wahrheiten gibt. Und von denen, die diese Möglichkeit erwähnen, nehmen viele sie nicht ernst. [...] Normative Begriffe bilden eine eigene, grundlegende Kategorie – wie etwa auch zeitliche oder logische Begriffe. [...] Mein Ziel

wird es sein [...], ihr distinktes Merkmal herauszuarbeiten: ihre Normativität. Eine Möglichkeit, dass zu tun, besteht in der Beschreibung von Fällen, in denen Normativität ganz offenkundig zu Tage tritt. Das ist leicht getan. Zwei solcher Fälle sind die Schlechtigkeit von Leiden sowie der Grund, den jemand hat, wenn er aus einem brennenden Gebäude springt. Aber Beispiele können missverständlich sein. Normativität kann mit anderen Merkmalen dieser Fälle verwechselt werden.<sup>88</sup>

Derlei »Verwechslungen« macht Parfit bei einer ganzen Reihe von Philosophen aus: Diese verwendeten zwar normative Wörter wie »Sollen«, »Grund« oder »rational«, würden damit aber häufig *psychische* Begriffe assoziieren. Damit träfen sie etwa Aussagen darüber, wie *psychisch zwingend* bestimmte Überlegungen oder andere mentale Ereignisse eine Person in ihrem Denken und Handeln beeinflussten. Dabei lasse sich in vielen Fällen die begründete Vermutung aufstellen, dass diese Philosophen die Eigenständigkeit des Normativen wenigstens unterschwellig oder im Alltag durchaus anerkannt haben. Parfit nennt David Hume, Bernard Williams oder John Mackie als Beispiele.

### 3.3 Die Ontologie des Normativen

Selbst wenn wir jedoch über das Normative Überlegungen anstellen können: Was garantiert, dass es derlei Tatsachen überhaupt gibt? – Parfit geht in der Beantwortung der Frage zweistufig vor: In einem *ersten Schritt* bemüht er sich, die Menge an Dingen, die das fundamental Normative ausmachen, möglichst gering zu halten. In einem *zweiten Schritt* stellt er diese Dinge durch Vergleich mit nicht-normativen Tatsachen der Logik und Mathematik als ontologisch »gewichtlos« und damit unschuldig dar.

Im ersten Schritt soll durch die Verlagerung des normativen Fundaments von Wert-Tatsachen auf Tatsachen über Gründe die Ausgangsstellung in ontologischen Fragen gegenüber dezidiert »naturalistischen« Positionen gerade im Vergleich zum nicht-naturalistischen Realisten verbessert werden. Nicht allein wird durch die

<sup>88</sup> Parfit 2006, S. 330f. (Übers.: Hg.).

Verminderung der Arten fundamental normativer Tatsachen die entsprechende Begründungslast *quantitativ* verringert, auch *qualitativ* scheint die Existenz von Gründe-Tatsachen gegenüber Wert-Tatsachen unbedenklicher zu sein: Dass bestimmten Handlungen, Charaktertypen und anderen Dingen »Werteigenschaften« zukommen sollen, die nicht lediglich Projektionen unserer persönlichen Einstellungen sind, wirkt *prima facie* »merkwürdiger« als die Rede davon, dass bestimmte Tatsachen für oder gegen etwas sprechen.

Bezüglich des *quantitativen Vorteils* lässt der Bezug auf die im Vergleich zu konkreten Einzeldingen und ihren Eigenschaften weniger greifbaren Tatsachen es intuitiv offener erscheinen, ob diese nicht als Träger normativer Eigenschaften des Dafür- oder Dagegensprechens in Frage kommen könnten. Neben derlei Tatsachen als den Trägern dieser Eigenschaften bleiben auch die Eigenschaften selbst sogar im Vergleich zu sogenannten »dünnen«, deskriptiv unbestimmten Werteigenschaften wie Gutsein oder Schönsein philosophisch abstrakter und verhalten sich damit naturalistischen Intuitionen gegenüber flexibler.

Bezüglich des *qualitativen Vorteils* wird der von anderen anti-realistischen Positionen gefürchtete ontologische »Ballast« durch den Kognitivismus gemindert, dem zufolge normative Tatsachen durch *nicht*-normative Tatsachen konstituiert werden: Die Tatsache, dass eine bestimmte *andere Tatsache* – etwa die Tatsache, dass das Messer scharf ist – für oder gegen ein bestimmtes Verhalten spricht – hier also: damit vorsichtig zu hantieren –, ist Merkmalen dieser anderen Tatsache geschuldet, und diese Merkmale sind über ontologisch unproblematische (weil nicht-normative) Eigenschaften von Ereignissen und Dingen bestimmbar. Dessen unbeschadet drückt erst die Tatsache, dass diese Eigenschaften von Ereignissen oder Zuständen für oder gegen ein Verhalten sprechen, den nicht-normativen Tatsachen einen eigenständigen normativen Stempel auf.

An dieser Stelle drängt sich eine Frage auf: Wenn die normative die nicht-normative Tatsache konstituiert, das heißt in nicht-kausaler Weise hervorbringt, warum können wir dann nicht annehmen, dass beide letztlich ein und dieselbe Tatsache sind? Im dritten Band von OWM geht Parfit noch einmal genauer<sup>89</sup> auf diese Frage ein

89 Siehe OWM, Bd. 2, Kap. 26, und OWM, Bd. 3, Kap. 39. Parfit verwendet für die Konstitutionsbeziehung den Begriff der »non-causal making relation«.

und betrachtet als Beispiel das Verhältnis der normativen Tatsache, dass eine Handlung moralisch richtig ist, zur nicht-normativen Tatsache, dass sie ein Leiden minimiert. Nehmen wir an, dass der Begriff moralischer Richtigkeit und der Begriff des minimalen Leidens, die für die Beschreibung dieser Tatsachen verwendet werden, auf Eigenschaften verweisen, die notwendigerweise koextensional sind, ebenso wie im Falle der Eigenschaften des H<sub>2</sub>O-Seins und des Wasser-Seins. Im letzteren Fall scheint nicht nur zu gelten, dass H<sub>2</sub>O-Sein das Wasser-Sein *ontologisch konstituiert*, sondern auch, dass beide Eigenschaften *miteinander identisch* sind. Wieso sollte man dann nicht dasselbe für die Eigenschaften des minimalen Leidens und die Eigenschaft moralischer Richtigkeit sagen können?

Parfit argumentiert, dass die Konstitutionsbeziehung auch bei miteinander identischen Eigenschaften bestehen *kann*, aber nicht jede derartige Beziehung auf die Identität beider verweist. Während H<sub>2</sub>O-Sein konstitutiv für Wasser-Sein ist *und* beide Eigenschaften miteinander identisch sind, gilt dasselbe beispielsweise *nicht* für das Eigenschaftspaar riskantes Fahren + Illegalität: Auch wenn mein riskantes Fahren konstitutiv für die Illegalität meines Tuns ist, sind das doch zwei verschiedene Eigenschaften. Ebenso verhält es sich nach Parfit mit normativen Eigenschaften und den sie wahrmachenden nicht-normativen Eigenschaften. Sie entstammen grundverschiedenen Kategorien:

Es gibt viele solcher nicht-überlappenden Kategorien. Physikalische Tatsachen können beispielsweise nicht dasselbe sein wie logische, juristische, musikalische, grammatikalische, exegetische oder mythologische Tatsachen. Auch könnte keine dieser anderen Tatsachen in zwei der genannten Kategorien fallen. Wie diese Beispiele nahelegen, wäre es nicht überraschend, wenn (wie ich glaube gezeigt zu haben) nicht-empirische Wahrheiten nicht dieselben sein können wie irgendwelche naturalistisch erklärbaren und empirisch auffindbaren Tatsachen.<sup>90</sup>

90 OWM, Bd. 3, S. 84 (Übers.: Hg.).

### 3.4 Die Epistemologie des Normativen

Während normative Tatsachen auf unsere Erkenntnisvermögen nicht kausal einwirken – sie *verursachen* unsere Einstellungen und Handlungen nicht –, können sie dessen unbeschadet eine Rolle bei der *Rechtfertigung* vieler unserer Überzeugungen und Handlungen spielen. Wiederum dienen Parfit die mathematischen Wahrheiten als vergleichbarer Fall: Ebenso wie wir mathematisch korrekt räsionieren können, auch ohne dass wir dabei in kausalem Kontakt mit Zahlen stehen, erkennen wir die Gültigkeit normativer Wahrheiten an und handeln danach.

Dass wir ihnen entsprechend handeln können, macht zwei weitere Eigenschaften mathematischer und normativer Wahrheiten deutlich: Zum einen sind nicht allein ihr *Einfluss auf uns*, sondern auch unsere *Reaktionen auf sie* nicht kausaler Natur. Auf beide Arten von Tatsachen können wir angemessen reagieren, indem wir richtig rechnen oder das tun, wofür wir Gründe haben. Das heißt zweitens jedoch *nicht*, dass die Erkenntnis mathematischer oder normative Wahrheiten mit *keinerlei* kausalen Reaktionen verbunden ist.

Im Falle mathematischer Wahrheiten verdeutlicht Parfit dies am Beispiel von Computern: In einem Computer finden während der Informationsverarbeitung eine Vielzahl von kausalen Prozessen statt, die natürlichen Gesetzmäßigkeiten gehorchen. Gleiches ist für uns Menschen anzunehmen, die bei der Erkenntnis einer moralischen Wahrheit mittels kausal beschreibbarer, neurophysiologischer Vorgänge in unserem Gehirn nicht allein bestimmte Überzeugungen bilden, sondern sie auch in der raumzeitlichen Wirklichkeit im Einklang mit den Naturgesetzen angemessen umzusetzen versuchen. Parfit sieht kausale und rationale Formen der Beeinflussung also nicht als einander widerstreitend an, sondern als miteinander kompatibel. So schreibt er für den Fall epistemischer Rationalität:

Wir bilden viele unserer Überzeugungen, indem wir in gültiger Weise Schlüsse ziehen, und Gültigkeit ist keine physikalische Eigenschaft. Aber diese nicht-physikalischen Erklärungen sind mit physikalischen Gesetzen *kompatibel*. Wir können verstehen, wie die physikalischen Prozesse in unse-

ren Gehirnen dem gültigen Schlussfolgern korrespondieren. Obgleich diese Fähigkeiten *nicht-kausale* Reaktionen auf die Gültigkeit solcher Überlegungen beinhalten, sind derartige Reaktionen nicht *anti-kausal* oder übernatürlich, wie die einfacheren Fähigkeiten von Computern zeigen.<sup>91</sup>

Die Behauptung der Eigenständigkeit unserer rationalen Fähigkeiten zur Erkenntnis von und Reaktion auf normative Wahrheiten muss sich gegen Ansätze zur Wehr setzen, die mit der Naturalisierbarkeit dieser Fähigkeiten deren Eigenständigkeit leugnen. Während Parfit die Naturalisierbarkeit mit Bezug auf ihre *Entwicklung* nicht bestreitet – er nimmt hier die darwinistische Evolutionstheorie in Anspruch –, verwirft er den Vorschlag naturalistischer Philosophen, die auch die Naturalisierbarkeit ihrer *Ausgestaltung* behaupten. Jenen Philosophen zufolge können wir unsere theoretischen und praktischen normativen Einstellungen sehr gut dadurch erklären, dass diese während unserer biologischen Entwicklung unserem Überleben dienlich waren. Damit erscheint es überflüssig, zusätzlich anzunehmen, dass wir so in die Lage versetzt wurden, normative Wahrheiten zu erfassen, für deren Geltung wir über keine unabhängige respektable und empirisch gestützte Methode verfügen.<sup>92</sup>

Wie andere Philosophen auch<sup>93</sup> stellt Parfit dieses Argument in Frage. Insbesondere bezweifelt er, dass alle unsere normativen Einstellungen nützlich für unser Überleben waren und dass wir für die Wahrheit wenigstens einiger normativer Überzeugungen tatsächlich keine guten Gründe haben. So wären die meisten unserer Überzeugungen auch dann für uns nützlich, wenn wir nicht der Überzeugung wären, gute epistemische Gründe dafür zu haben, sie zu glauben, sondern sie schlicht ausbilden würden. Gleichfalls wären die meisten unserer Handlungen auch dann für uns nützlich, wenn wir nicht der Überzeugung wären, gute praktische Gründe zu haben, die uns zu ihnen motivieren:<sup>94</sup> Parfit nennt als Beispiel

91 OWM, Bd. 2, S. 500 (Übers.: Hg.).

92 Eine einflussreiche Version eines solchen Arguments stammt von Sharon Street, auf das sich Parfit auch bezieht, vgl. Street 2006 und OWM, Bd. 2, S. 492, Kap. 33 sowie OWM, Bd. 3, S. 264 ff.

93 Vgl. etwa Lazari-Radek/Singer 2012.

94 Vgl. OWM, Bd. 2, Abschnitte 117 und 118.

die Überzeugung, dass uns die Schlechtigkeit von Schmerzen gute Gründe gibt, sie zu vermeiden. Weiterhin bieten viele normative Überzeugungen, etwa die Goldene Regel, keine Überlebensvorteile.<sup>95</sup>

Damit wir der gerechtfertigten Auffassung sein können, dass bestimmte normative Überzeugungen einfach deshalb gebildet wurden, weil sie wahr sind – Parfit stellt hier die Überlegung an, dass unsere Fähigkeit zu vernünftigem Denken und Handeln zwar in vielerlei Hinsicht vorteilhaft für unser Überleben war, sich dann aber »verselbstständigte« und die Tür zur Erkenntnis weiterer Wahrheiten aufgestoßen hat –, müssen wir ausschließen, dass unsere moralischen Überzeugungen nicht zu sehr divergieren und beispielsweise je nach kulturellem Hintergrund unterschiedlich ausfallen. Wie zu Beginn von Abschnitt 2 bereits hervorgehoben, könnten solche unauflösbaren Meinungsverschiedenheiten dafür sprechen, dass normative Sätze weder wahr noch falsch sind und es daher letztlich auf gar nichts ankommt. Parfit stellt deshalb die *Kovertenzbehauptung* [*Convergence Claim*] auf:

Würde jedermann Kenntnis von allen relevanten nicht-normativen Tatsachen haben, dieselben normativen Begriffe verwenden, die relevanten Argumente verstehen und sorgfältig über sie nachdenken und dabei nicht unter störenden Einflüssen stehen, dann hätten wir alle ähnliche normative Überzeugungen.<sup>96</sup>

Zwar sei es für den nicht-realistischen Kognitivisten schwierig, diese Behauptung zu beweisen, doch hätten normative Skeptiker – anders als man meinen könnte – keinen leichteren Stand. Um Letzteres zu zeigen, bemüht sich Parfit nun *erstens*, die vielen »Störfaktoren« herauszuarbeiten, die normative Meinungsverschiedenheiten erklären können: So bestünden bisweilen tatsächlich nur Uneinigkeiten über nicht-normative oder metanormative Sachverhalte. Die Streitparteien wären unter Umständen nicht voll informiert, würden unterschiedliche Begriffe verwenden oder seien epistemischen Verzerrungen ausgesetzt. Weiterhin seien viele normative Wahrheiten graduell oder unpräzise, und schließlich gäbe es

95 Vgl. ebd., S. 536.

96 Ebd., S. 546 (Übers.: Hg.).

auf einige unserer normativen Fragen auch schlicht keine Antwort. *Zweitens* hätten wir überzeugende Gründe, davon auszugehen, bereits über gute Antworten auf viele unserer normativen Fragen zu verfügen. Normative Sachverhalte wie diejenigen, dass Schmerzen fast immer schlecht sind oder dass unschuldige Menschen nicht leichtfertig getötet werden dürfen, könne man nicht (mehr) ernsthaft in Zweifel ziehen.

### 3.5 Die Rezeption der Metaethik

Die in den zweiten Band von OWM aufgenommenen Kommentare zu Parfits Ideen waren noch ausschließlich auf seine normativ-ethischen Thesen fokussiert – ein Umstand, der sich – wie eingangs beschrieben – daraus ergibt, dass diese Kommentare größtenteils auf den *Tanner Lectures* von 2002/03 basieren: Gegenstand dieser Vorlesungen waren Thesen und Argumente, in denen Parfit seine *Triple Theory* entwickelt.

Demgegenüber hat sich spätestens seit der Publikation von OWM eine lebhafte Debatte um Themen und Thesen von Parfits Metaethik entwickelt. In Buchform erschienen sind einige dieser Beiträge bereits 2009<sup>97</sup> in Jussi Suikkanens *Essays on Derek Parfit's On What Matters*, darunter zwei Beiträge von James Lenman und Michael Smith, die sich mit Parfits Kritik des Naturalismus und der wunschbasierten Theorie der Gründe auseinandersetzen und sich bemühen, beide Positionen gegen Parfit zu verteidigen.<sup>98</sup>

Derzeit erscheinen zusammen mit dem dritten Band von OWM gleich mehrere Werke, die hauptsächlich oder zumindest auch Kritiken und Anmerkungen zu Parfits Metaethik enthalten: *Does Anything Really Matter?*, herausgegeben von Peter Singer;<sup>99</sup> *Reading Parfit*, herausgegeben von Simon Kirchin;<sup>100</sup> und (für den deutschsprachigen Raum) *Worauf es ankommt*, veröffentlicht von den Herausgebern dieses Bandes.<sup>101</sup> Zu allen drei Veröffentlichungen hat Parfit Erwidern verfasst. Neben diesen Büchern gibt es eine

97 Dabei sind die Beiträge minus die Einleitung von Suikkanen im selben Jahr bereits als *Special Issue* der Zeitschrift *Ratio* veröffentlicht worden.

98 Vgl. Lenman 2009 und Smith 2009.

99 Vgl. Singer 2017.

100 Vgl. Kirchin (i. E.).

101 Vgl. Hoesch/Muders/Rüther (i. E.).

Vielzahl von Fachaufsätzen, die sich mit seiner Metaethik auseinandersetzen und etwa seinen »dritten Weg« zwischen (nicht-naturalistischem) Realismus und Anti-Realismus diskutieren.<sup>102</sup>

Der Band von Singer verdient hierbei besondere Erwähnung, da Parfitts Erwiderungen darauf zusammen mit weiteren Überlegungen den dritten Band von OWM ausmachen. Da die dort versammelten Kritiken sichtbaren Einfluss auf die Weiterentwicklung seiner Metaethik genommen haben, sei Parfitts dortiger Umgang mit seinen Kritikern kurz beleuchtet.

Auffallend ist ein deutlicher Strategiewechsel. Setzte der zweite Band von OWM in der Metaethik eher auf Ausgrenzung statt auf Inklusion, also auf die Verteidigung der Eigenständigkeit seines nicht-realistischen Kognitivismus als vielversprechendste Option, bemüht sich Parfitt im dritten Band von OWM, die Anschlussfähigkeit seiner Theorie aufzuzeigen. Um in dem Bild zu bleiben, das Parfitt zur Charakterisierung seines Projektes einer Vereinigung verschiedener Moraltheorien bemüht (siehe dazu Abschnitt 2.4), erhält der gemeinsam zu bezwingende Berg nun auch in der Metaethik einen eigenen Gipfel. Sowohl auf Seiten der naturalistischen Realisten als auch unter den Expressivisten sucht (und findet) Parfitt nun Verbündete. Während er seinen metaethischen Mitbewerbern auch bereits in den ersten Bänden von OWM Vorschläge zur Modifikation ihrer Theorien machte, tendiert er nun zusätzlich dazu, einige seiner vormaligen Deutungen konkurrierender Metaethiken zu deren Gunsten abzuändern. Im Ergebnis bedeutet das im Regelfall, dass er dort nun Raum für ontologiebefreite normative Wahrheiten von der Art findet, wie er sie selbst verteidigt; so identifiziert er naturalistische und expressivistische Varianten des nicht-realistischen Kognitivismus.<sup>103</sup> Zwei von Parfitts Kollegen – Allan Gibbard und Peter Railton – nähern sich mit neuerlichen Kommentaren auf Parfitts Reaktionen nun umgekehrt Parfitts Position weiter an.<sup>104</sup>

Mit alledem ist freilich nicht gesagt, dass keine tiefgehenden metaethischen Divergenzen verbleiben. Viele der von Parfitts Kritikern

102 Vgl. z. B. Suikkanen 2016 und Rütter 2016. Zu diesem Thema vgl. aber auch Muders/Rütter (i. E.).

103 Vgl. Railton 2017; Gibbard 2017; OWM, Bd. 3, Kapitel. 42 und 46f. Der naturalistischen Variante seiner Position gibt er den Namen »erweiterter Naturalismus« [*wider Naturalism*].

104 Abgedruckt als Abschnitte 143 und 155 in OWM, Bd. 3.

vorgebrachten Einwände, die sich auf die Fundamentalität, Ontologie und Epistemologie des Normativen beziehen, werden von Parfit abgewiesen. Wie seine Er widerungen u. a. auf Smith, Blackburn, Street und andere zeigen, erscheint die Metaethik auch aus Parfits Werte unverändert zerklüftet, und das von Gibbard ausgemachte »Happy Ending«<sup>105</sup> markiert bestenfalls einen Nebengipfel auf dem Weg nach oben. Dennoch fühlt sich Parfit ermutigt, von einer neuen *Triple Theory* zu sprechen, die neue Pfade zum Gipfel weisen kann:

Obwohl Railton, Gibbard und ich früher die zentralen Behauptungen der jeweils anderen zurückgewiesen haben, ist das nicht länger der Fall. Unsere drei Sichtweisen haben sich einander angeglichen. [...] [D]iese Sichtweisen unterscheiden sich zwar [...] immer noch in verschiedenen Hinsichten, und wir verfolgen teilweise unterschiedliche Ziele, doch es gibt keine fundamentalen Differenzen. Wir hoffen, dass andere zu ähnlichen Schlüssen kommen werden.<sup>106</sup>

\*\*\*

Es sollte deutlich geworden sein, dass es nicht möglich ist, die Philosophie Derek Parfits als Entfaltung einer einheitlichen, sich durch das Werk ziehenden Grundidee zu beschreiben: Parfit lässt sich keiner philosophischen Schule zuordnen, knüpft an ganz verschiedene philosophische Themen und Traditionen an und vermeidet allzu enge systematische »Bande« zwischen den jeweiligen Teilen seiner Theorie. Will man daher unsachliche Narrative oder Etikettierungen vermeiden, bleibt als schlankes Fazit übrig, dass wir in Parfit einen eigenständigen Denker vor uns haben, mit dessen Aufsätzen und Büchern Fragen ins Zentrum der Philosophie des Geistes und der Ethik geraten sind, die die analytische Philosophie teils vergessen und teils noch gar nicht entdeckt hatte.

Wie viele große Denker musste aber auch Parfit einiges ungesagt lassen. Dieses Manko betrifft, wie er am Ende des dritten Bandes

105 So der Titel von Abschnitt 156 ebd.

106 Ebd., S. 236 (Übers.: Hg.).

von OWM selbst festhält, sogar die Titelfrage seiner letzten großen Unternehmung, also die Frage, worauf es denn nun letztlich ankommt:

Ich bedaure, in einem Buch mit dem Titel *On What Matters* so wenig davon gesprochen zu haben, worauf es ankommt. In meinem dann vierten Band werde ich hoffentlich mehr dazu sagen können.<sup>107</sup>

Zwei Punkte spricht Parfit aber immerhin an. Der erste thematisiert das vermeidbare Leid der Gegenwart:

Eine Sache, auf die es besonders ankommt, ist das Versagen von uns Reichen. Obgleich es so einfach wäre, unterlassen wir es, einen Großteil des Leids und der vielen frühzeitigen Tode der ärmsten Menschen auf der Welt zu verhindern.<sup>108</sup>

Noch wichtiger erscheint ihm aber das langfristige Überleben der Menschheit:

Am meisten kommt es jetzt darauf an, wie wir auf die verschiedenen Gefahren reagieren, die das Überleben der Menschheit bedrohen. Einige dieser Gefahren sind selbstgemacht, und wir finden gerade immer mehr darüber heraus, wie wir auf sie, aber auch auf andere Gefahren reagieren können. Wenn wir diese Risiken vermindern und die Menschheit die kommenden Jahrhunderte übersteht, könnten unsere Nachkommen oder Nachfolger diese Gefahren endgültig abwenden, indem sie sich über unsere Galaxie ausbreiten.<sup>109</sup>

Bedauerlicherweise konnte Parfit diese Überlegungen nicht mehr fortführen; die Dringlichkeit beider Projekte wird aber von vielen Ethikern erkannt. Man darf daher optimistisch sein, dass weder die inhaltliche Arbeit an Parfits Themen noch die Debatte um seine ethische Theorie versiegen werden.

107 Ebd., S. 436 (Übers.: Hg.).

108 Ebd. (Übers.: Hg.).

109 Ebd. (Übers.: Hg.).

## Literatur

- Arrhenius, Gustaf/Ryberg, Jesper/Tännsjö, Torbjörn 2014: »The Repugnant Conclusion.« In: Edward N. Zalta (Hg.): *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), (<http://plato.stanford.edu/archives/spr2014/entries/repugnant-conclusion/>).
- Arrhenius, Gustaf 2016: »Population Ethics and Different-Number-Based Imprecision.« *Theoria* 82 (2), S. 166-181.
- Attfield, Robin 2014: *Environmental Ethics. An Overview for the Twenty-First Century*, Cambridge.
- Bailey, Andrew M. 2016: »You Are an Animal.« *Res Philosophica* 93 (1), S. 205-218.
- Blackburn, Simon 1993: *Essays in Quasi-Realism*, New York.
- Blackorby, Charles/Walter Bossert/David Donaldson (Hg.) 2005: *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*, Cambridge.
- Broome, John 1992: *Counting the Cost of Global Warming*, Cambridge.
- Chang, Ruth 1997: »Introduction.« In: Dies. (Hg.): *Incommensurability, Incomparability and Practical Reason*, Cambridge, S. 1-34.
- Chang, Ruth 2002: »The Possibility of Parity.« *Ethics* 112 (4), S. 659-688.
- Chang, Ruth 2016: »Parity, Imprecise Comparability, and the Repugnant Conclusion.« *Theoria* 82, S. 183-215.
- Cohen, A. I. 2009: »Compensation for Historic Injustices. Completing the Boxill and Sher Argument.« *Philosophy & Public Affairs* 37 (1), 81-102.
- Cummiskey, David 1990: »Kantian Consequentialism.« *Ethics* 100 (3), S. 586-615.
- Cummiskey, David 1996: *Kantian Consequentialism*, New York/Oxford.
- Dancy, Jonathan (Hg.) 1997: *Reading Parfit*, Oxford.
- Darwall, Stephen 2014: »Agreement Matters. Critical Notice of Derek Parfit, *On What Matters*.« *Philosophical Review* 123 (1), S. 79-105.
- Dath, Dietmar 2017: »Ethik für Monster.« *Frankfurter Allgemeine Zeitung*, 11. 1. 2017, S. N3.
- de Lazari-Radek, Katarzyna/Peter Singer 2012: »The Objectivity of Ethics and the Unity of Practical Reason.« *Ethics* 123 (1), S. 9-31.
- Dufner, Annette (i. E.): »Prioritaristischer Konsequentialismus?« In: Matthias Hoesch/Sebastian Muders/Markus Rüter (Hg.): *Worauf es ankommt. Derek Parfit in der Diskussion*, Hamburg.
- Dworkin, Ronald 1981: »What is Equality? Part 2: Equality of Resources.« *Philosophy and Public Affairs* 10 (4), S. 283-345.

- Dworkin, Ronald 2012: *Gerechtigkeit für Igel*, Berlin.
- Frankfurt, Harry G. 1988 [1971]: »Freedom of the Will and the Concept of a Person.« In: Ders.: *The Importance of What We Care About: Philosophical Essays*, Cambridge, S. 11-25.
- Fried, Barbara H. 2012: »What Does Matter? The Case for Killing the Trolley Problem (Or Letting It Die).« *Philosophical Quarterly* 62 (248), S. 505-529.
- Gesang, Bernward 2011: *Klimaethik*, Berlin.
- Gibbard, Allan 2017: »Parfit on Normative Concepts and Disagreement.« In: Peter Singer (Hg.): *Does Anything Really Matter? Parfit on Objectivity*, Oxford, S. 61-80.
- Goldin, Ian (Hg.) 2014: *Is the Planet full?*, Oxford.
- Hare, R. M. 1993: »Could Kant Have been a Utilitarian?« *Utilitas* 5 (1), S. 1-16.
- Herman, Barbara 1989: »Murder and Mayhem.« *The Monist* 72 (3), S. 411-431.
- Herman, Barbara 2011: »A Mismatch of Methods.« In: Derek Parfit: *On What Matters*, Bd. 2. Hg. von Samuel Scheffler, Oxford, S. 83-115.
- Herrmann, Martina 2001: »Der Personbegriff in der analytischen Philosophie.« In: Dieter Sturma (Hg.): *Person. Philosophiegeschichte – Theoretische Philosophie – Praktische Philosophie*, Paderborn, S. 167-189.
- Hoesch, Matthias/Sebastian Muders/Markus Rüter (Hg.) (i. E.): *Worauf es ankommt. Derek Parfit in der Diskussion*, Hamburg.
- Hoesch, Matthias/Martin Sticker (i. E.): »Parfit über Kantianismus und Konsequentialismus.« In: Matthias Hoesch/Sebastian Muders/Markus Rüter (Hg.): *Worauf es ankommt. Derek Parfit in der Diskussion*, Hamburg.
- Hooker, Brad 2011: »Rule Consequentialism.« In: Edward N. Zalta (Hg.): *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), (<http://plato.stanford.edu/entries/consequentialism-rule/>).
- Johnston, Marc 2003: »Human Concerns without Superlative Selves.« In: Raymond Martin/John Barresi (Hg.): *Personal Identity*, Oxford, S. 260-291.
- Kant, Immanuel: *Grundlegung zur Metaphysik der Sitten*, in: *Gesammelte Schriften* Bd. 4. Hg. von der Preußischen Akademie der Wissenschaften, Berlin 1900 ff. (zitiert als *GMS*).
- Kirchin, Simon (Hg.) (i. E.): *Reading Parfit: On What Matters*, London.

- Krebs, Angelika 2000: »Einleitung: Die neue Egalitarismus-Kritik im Überblick.« In: Dies. (Hg.): *Gleichheit oder Gerechtigkeit. Texte der neuen Egalitarismuskritik*, Frankfurt am Main, S. 7-37.
- Lenman, James 2009: »Naturalism without Tears.« In: Jussi Suikkanen/John Cottingham (Hg.): *Essays on Derek Parfit's On What Matters*, Malden (Mass.), S. 1-20.
- Locke, John 1981 [1689]: *Versuch über den menschlichen Verstand*, Band 1. Hg. v. Reinhard Brandt, Hamburg.
- MacFarquhar, Larissa 2011: »How to Be Good.« *The New Yorker*, Annals of Ideas, 2011 Issue (September), online unter <http://www.newyorker.com/magazine/2011/09/05/how-to-be-good>.
- McGee, Andrew 2016: »We Are Human Beings.« *Journal of Medicine and Philosophy* 41 (2), S. 148-171.
- Meyer, Lukas H. 2006: »Nach dem Recht auf Rückkehr.« *Deutsche Zeitschrift für Philosophie* 54 (4), S. 623-639.
- Muders, Sebastian/Markus Rüter (i. E.): »Tatsachen ohne Metaphysik?« In: Matthias Hoesch/Sebastian Muders/Markus Rüter (Hg.): *Woraufes ankommt. Derek Parfit in der Diskussion*, Hamburg.
- Nagel, Thomas 2013: *Geist und Kosmos. Warum die materialistische neodarwinistische Konzeption der Natur so gut wie sicher falsch ist*, Berlin.
- Olson, Eric T. 1997: *The Human Animal. Personal Identity without Psychology*, Oxford.
- Olson, Eric T. 2002: »Personal Identity.« In: Edward N. Zalta (Hg.): *The Stanford Encyclopedia of Philosophy* (Fall 2002 Edition), (<https://plato.stanford.edu/archives/fall2002/entries/identity-personal/>).
- Olson, Eric T. 2015: »On Parfit's View That We Are Not Human Beings.« *Royal Institute of Philosophy Supplement* 76, S. 39-56.
- Otsuka, Michael 2009: »The Kantian Argument for Consequentialism.« In: Jussi Suikkanen/John Cottingham (Hg.): *Essays on Derek Parfit's On What Matters*, Malden (Mass.), S. 61-78.
- Otsuka, Michael 2012: »Prioritarianism and the Separateness of Persons.« *Utilitas* 24 (3), S. 365-380.
- Parfit, Derek 1976: »On Doing the Best for Our Children.« In: Michael D. Bayles (Hg.): *Ethics and Population*, Cambridge (Mass.), S. 100-115.
- Parfit, Derek 1982: »Future Generations. Further Problems.« *Philosophy and Public Affairs* 11, S. 113-172.
- Parfit, Derek 1983: »Energy Policy and the Further Future. The Social Dis-

- count Rate.« In: Douglas MacLean/Peter Brown (Hg.): *Energy and the Future*, Totowa (NJ), S. 166-179.
- Parfit, Derek 1984: *Reasons and Persons*, Oxford (zitiert als RaP).
- Parfit, Derek 1986: »Overpopulation and the Quality of Life.« In: Peter Singer (Hg.): *Applied Ethics*, Oxford, S. 145-164. Wiederabgedruckt in Jesper Ryberg/Torbjörn Tännsjö (Hg.) 2004: *The Repugnant Conclusion*, Dordrecht, S. 7-22.
- Parfit, Derek 1997a: »Equality and Priority.« *Ratio* 10, S. 202-221.
- Parfit, Derek 1997b: »Reasons and Motivation.« *Proceedings of the Aristotelian Society*, Supplementary Volume 71, S. 99-130.
- Parfit, Derek 2001: »Rationality and Reasons.« In: Dan Egonsson/Jonas Josefsson/Björn Petterson/Toni Rønnow-Rasmussen (Hg.): *Exploring Practical Philosophy. From Action to Values*, Aldershot, S. 17-39.
- Parfit, Derek 2002 [1991]: »Equality or Priority?« In: Matthew Clayton/Andrew Williams (Hg.): *The Ideal of Equality*, Basingstoke, S. 81-125. Zuvor im Jahr 1995 veröffentlicht als *The Lindley Lecture 1991*, Lawrence.
- Parfit, Derek 2004a: »Postscript.« In: Jesper Ryberg/Torbjörn Tännsjö (Hg.): *The Repugnant Conclusion*, Dordrecht, S. 257.
- Parfit, Derek 2004b: »What We Could Rationally Will.« In: Grethe Peterson (Hg.): *The Tanner Lectures on Human Values XXIV*, Salt Lake City, S. 285-369.
- Parfit, Derek 2006: »Normativity.« In: Russ Shafer-Landau (Hg.): *Oxford Studies in Metaethics*, Bd. 1, Oxford, S. 325-380.
- Parfit, Derek 2011: *On What Matters*, Bde. 1 und 2. Hg. von Samuel Scheffler, Oxford (zitiert als OWM, Bd. 1, und OWM, Bd. 2).
- Parfit, Derek 2012: »Another Defence of the Priority View.« *Utilitas* 24 (3), S. 399-440.
- Parfit, Derek 2013: »Worauf es wirklich ankommt«. Ein Interview mit Derek Parfit.« *Hohe Luft* 2013 (2), S. 20-33.
- Parfit, Derek 2017: *On What Matters*, Bd. 3, Oxford (zitiert als OWM, Bd. 3).
- Persson, Ingmar 2012: »Prioritarianism and Welfare Reductions.« *Journal of Applied Philosophy* 29 (4), S. 289-301.
- Quante, Michael (Hg.) 1999: *Personale Identität*, Stuttgart.
- Quante, Michael 2002: *Personales Leben und menschlicher Tod*, Frankfurt am Main.
- Quante, Michael 2007: *Person*, Berlin.
- Railton, Peter 2017: »Two Sides of the Meta-Ethical Mountain?« In: Peter Singer (Hg.): *Does Anything Really Matter? Parfit on Objectivity*, Oxford, S. 35-60.
- Rawls, John 1971: *A Theory of Justice*, Cambridge (Mass.)/London.

- Roberts, Melinda A. 2011: »An Asymmetry in the Ethics of Procreation.« *Philosophy Compass* 6 (11), S. 765-776.
- Roberts, Melinda A. 2015: »The Nonidentity Problem.« In: Edward N. Zalta (Hg.): *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition), (<http://plato.stanford.edu/archives/win2015/entries/nonidentity-problem/>).
- Roberts, Melinda A./David T. Wasserman (Hg.) 2009: *Harming Future Persons. Ethics, Genetics and the Nonidentity Problem*, New York.
- Roemer, John 1998: *Equality of Opportunity*, Cambridge (Mass.).
- Roser, Dominic/Christian Seidel 2015: *Ethik des Klimawandels. Eine Einführung*, Darmstadt.
- Ross, Jacob 2009: »Should Kantians Be Consequentialists?« In: Jussi Suikkanen/John Cottingham (Hg.): *Essays on Derek Parfit's On What Matters*, Malden (Mass.), S. 144-153.
- Ross, William David 2002 [1930]: *The Right and the Good*. Hg. von Philip Stratton-Lake, Oxford.
- Rüther, Markus 2016: »Normative Wahrheiten ohne Ontologie? Derek Parfit und der neue Non-Naturalismus.« *Zeitschrift für Praktische Philosophie* 3 (2), S. 187-220.
- Scanlon, Thomas 1998: *What We Owe to Each Other*, Cambridge (Mass.).
- Scheffler, Samuel 2011: »Introduction.« In: Derek Parfit: *On What Matters*, Bd. I. Hg. von Samuel Scheffler, Oxford, S. xix-xxxii.
- Scruton, Roger 2014: »Parfit the Perfectionist.« *Journal of the Royal Institute of Philosophy* 89, S. 621-634.
- Singer, Peter 2011: »One Mountain.« *The Times Literary Supplement*, 20. Mai 2011.
- Singer, Peter (Hg.) 2017: *Does Anything Really Matter? Essays on Parfit on Objectivity*, Oxford.
- Smith, Michael 2009: »Desires, Values, Reasons, and the Dualism of Practical Reason.« In: Jussi Suikkanen/John Cottingham (Hg.): *Essays on Derek Parfit's On What Matters*, Malden (Mass.), S. 116-143.
- Snowdon, Paul 1997: »Persons and Personal Identity.« In: Sabina Lovibond/S. G. Williams (Hg.): *Identity, Truth, Value*, Oxford, S. 33-48.
- Sosa, Ernest 2003: »Surviving Matters.« In: Raymond Martin/John Barresi (Hg.): *Personal Identity*, Oxford, S. 199-215.
- Strawson, Peter 1959: *Individuals. An Essay in Descriptive Metaphysics*, Oxford.
- Street, Sharon 2006: »A Darwinian Dilemma for Realist Theories of Value.« *Philosophical Studies* 127 (1), S. 109-166.
- Sturma, Dieter 1997: *Philosophie der Person. Die Selbstverhältnisse von Subjektivität und Moralität*, Paderborn u. a.

- Sturma, Dieter (Hg.) 2001: *Person – Philosophiegeschichte – Theoretische Philosophie – Praktische Philosophie*, Paderborn.
- Suikkanen, Jussi (2016): »Non-Realist Cognitivism, Truth and Objectivity.« *Acta Analytica*, (<http://link.springer.com/article/10.1007/s12136-016-0300-5>).
- Suikkanen, Jussi/John Cottingham (Hg.) 2009: *Essays on Derek Parfit's On What Matters*, Malden (Mass.).
- Wilkes, Kathleen 1988: *Real People. Personal Identity without Thought Experiments*, Oxford.
- Wood, Allen 2011: »Humanity as End in Itself.« In: Derek Parfit: *On What Matters*, Bd. 2. Hg. von Samuel Scheffler, Oxford, S. 58-82.