



Inkompatibilismus und die absolutistische Konzeption von Vernunft

Prof. Dr. Marcus Willaschek

Exzellenzcluster
Die Herausbildung normativer Ordnungen

www.normativeorders.net

Goethe-Universität Frankfurt

Institut für Philosophie

Grüneburgplatz 1

60629 Frankfurt am Main

Willascheck@em.uni-frankfurt.de

Erschienen in:

Philosophisches Jahrbuch der Görres-Gesellschaft 115 (2008), 397-417



Dieses Werk ist unter einem Creative Commons Namensnennung-Keine kommerzielle Nutzung-Keine Bearbeitung 3.0 Deutschland Lizenzvertrag lizenziert. Um die Lizenz anzusehen, gehen Sie bitte zu <http://creativecommons.org/licenses/by-nc-nd/3.0/de/>.

1. Einleitung

Die Begriffe *Vernunft* und *Freiheit* verweisen wechselseitig aufeinander. Zwar gibt es Formen von Freiheit, über die auch die vernunftlose Natur verfügen kann, etwa ein gefangenes Tier, das entkommt und von dem wir sagen können, nun sei es wieder *frei*. Doch in einem anspruchsvolleren Sinn kann nur frei sein, wer über Vernunft verfügt. Dabei verstehe ich unter Vernunft ein Bündel von kognitiven Fähigkeiten, zu denen an erster Stelle gehört, Widersprüche im Denken und Handeln zu vermeiden sowie seine theoretischen wie praktischen Einstellungen kritisch hinterfragen und auf berechnete Nachfragen hin begründen zu können. Diese Fähigkeit zur Distanzierung von den eigenen Einstellungen und zu ihrer Überprüfung anhand von normativen Maßstäben ist ein zentrales Element menschlicher Freiheit. Insofern gilt: keine Freiheit ohne Vernunft.

Umgekehrt gibt es keine Vernunft ohne Freiheit. Zwar ist vernünftiges Denken häufig durch eine Art von Notwendigkeit und Alternativlosigkeit gekennzeichnet: Wer vernünftig ist und weiß, dass aus A B folgt, für den gilt: Wer A sagt, *muss* auch B sagen. Doch in dieser Art von Vernunftnotwendigkeit werden nur diejenigen eine Einschränkung unserer Freiheit sehen, die Freiheit mit Beliebigkeit verwechseln. Ein Denken, das von momentanen Bedürfnissen, zufälligen Assoziationen und sachfremden Einflüssen bestimmt wird, ist nicht frei, sondern fremdbestimmt. Nur die Orientierung an normativen Standards wie denen der Widerspruchsfreiheit und Begründbarkeit verhindert, dass unsere Gedanken zum Spielball kontingenter Kräfte werden. Und was für das vernünftige *Denken* und *Sprechen* gilt, das gilt *mutatis mutandis* auch für das vernünftige *Wollen* und *Handeln*. In diesem Sinn setzt Vernunft stets Freiheit voraus.

Im Folgenden möchte ich diesem wechselseitigen Zusammenhang zwischen Vernunft und Freiheit etwas weiter nachgehen, indem ich zu zeigen versuche, dass unterschiedliche Konzeptionen von Vernunft auch unterschiedliche Konzeptionen von Freiheit nach sich ziehen. Im nächsten Teil dieses Beitrags werde ich die dabei leitende Fragestellung erläutern: Sind Willensfreiheit und Determinismus miteinander vereinbar (2)? Anschließend geht es um zwei unterschiedliche Argumente für den Inkompatibilismus (also gegen die Vereinbarkeit von Freiheit und Determinismus), von denen sich herausstellen wird, dass sie auf strukturgleichen Hintergrundannahmen beruhen (3). Diese Hintergrundannahmen erweisen sich als Spezialfälle jenes Grundsatzes vernünftigen Denkens, den Kant als das „oberste

Prinzip der reinen Vernunft“ bezeichnet (4). Mit dem Kant der *Kritik der praktischen Vernunft* werde ich dann gegen den Kant der *Kritik der reinen Vernunft* argumentieren, dass dieses Prinzip keineswegs für die Vernunft als solche charakteristisch ist (5). Es stellt sich somit heraus, dass die inkompatibilistischen Argumente auf einer bestimmten („absolutistischen“) Konzeption von Vernunft beruhen, die keineswegs alternativlos ist. Es geht mir dabei nicht darum, den Inkompatibilismus zurückzuweisen, sondern nur darum, seine scheinbar „intuitive“ Plausibilität in Frage zu stellen, indem ich deutlich zu machen versuche, dass sie auf Hintergrundannahmen beruht, die ihrerseits nicht theoretisch unschuldig, sondern im Gegenteil höchst voraussetzungsreich sind. Abschließend werde ich in aller Kürze jene Alternative zu einer inkompatibilistischen Konzeption von Freiheit skizzieren, die durch die Ablehnung einer „absolutistischen“ Konzeption nahegelegt wird (6).

2. Willensfreiheit und Determinismus

Seit der Antike wird die Frage diskutiert, ob die Freiheit des Willens mit einer durchgängigen Notwendigkeit vereinbar ist, sei es, wie bei den Stoikern, mit der Notwendigkeit eines allumfassenden Schicksals, sei es, wie in der christlichen Philosophie des Mittelalters, mit der notwendigen Bestimmtheit aller Ereignisse durch den Willen Gottes, oder, wie in der Neuzeit, durch einen umfassenden physikalischen Determinismus. Erst kürzlich hat diese Frage wieder die Gemüter erregt, nachdem Neurowissenschaftler wie Wolf Singer und Gerhard Roth die These vertreten haben, unser Gehirn sei ein deterministisches System und Willensfreiheit daher eine Illusion.¹ Doch selbst wenn man die keineswegs unstrittige These des neuronalen Determinismus zugesteht, bleibt aus philosophischer Sicht die Frage bestehen, ob daraus tatsächlich folgt, dass unsere Entscheidungen nicht frei sind.

Auf diese Frage gibt es zwei mögliche Antworten, die man als Kompatibilismus und Inkompatibilismus bezeichnet. Kompatibilisten behaupten, dass eine Entscheidung auch dann frei sein kann, wenn sie vollständig durch Ursachen, die außerhalb der Kontrolle des Handelnden liegen, determiniert ist. Inkompatibilisten bestreiten dies. Dabei ist jener anspruchsvolle Sinn von Freiheit gemeint, in dem Freiheit eine notwendige Bedingung für Verantwortung ist. Unter Determinismus wird in diesem Zusammenhang die These verstanden, dass die Naturgesetze und der jeweils gegenwärtige Zustand der Welt nur eine einzige mögliche Zukunft zulassen. Genau das ist nach Meinung der Inkompatibilisten mit Freiheit unvereinbar: Eine Entscheidung ist nur dann frei, wenn sich die handelnde Person

¹ Vgl. dazu die in Geyer (2004) gesammelten Beiträge. Einen philosophisch anspruchsvollen und sehr hilfreichen Überblick über den aktuellen Stand der Willensfreiheitsdebatte gibt Keil (2007).

zwischen verschiedenen Möglichkeiten entscheiden kann – im einfachsten Fall zwischen dem Tun und dem Unterlassen einer bestimmten Handlung –, so dass man im Nachhinein sagen kann, die Person hätte auch anders handeln können. Doch wenn der Determinismus wahr ist, gibt es letztlich keine verschiedenen Möglichkeiten, zwischen denen man sich entscheiden kann. Die Naturgesetze und der gegenwärtige Zustand der Welt legen eindeutig fest, was die Person tun oder unterlassen wird. Niemand hätte jemals etwas anderes tun können als er tatsächlich getan hat. Freiheit und Determinismus sind demnach nicht miteinander vereinbar.

Der Inkompatibilismus scheint den gesunden Menschenverstand auf seiner Seite zu haben. Dennoch haben Philosophen seit der Antike zu zeigen versucht, dass er auf einem falschen Verständnis von Willensfreiheit beruht. Freiheit, so die *kompatibilistische* Auffassung, erfordert in erster Linie die Abwesenheit von innerem und äußerem Zwang. Nur wenn man den Determinismus mit einer Art von Zwang verwechselt, kann er als eine Gefahr für die Freiheit des Willens erscheinen. Auch verhindert der Determinismus den Kompatibilisten zufolge nicht, dass man anders hätte handeln können. Man hätte anders gehandelt, wenn man etwas anderes gewollt hätte.

Doch, so fragen die Inkompatibilisten zurück, hätte man auch etwas anderes *wollen* können? Für Freiheit reicht es ihrer Meinung nach nicht aus, dass man unter *anderen* Umständen etwas anderes getan hätte, sofern man keinen Einfluss darauf hatte, ob diese Umstände vorlagen oder nicht. Es komme vielmehr darauf an, dass unter exakt denselben Umständen, unter denen man A getan hat, man A auch hätte unterlassen können, so dass man sagen kann, dass es allein von der handelnden Person abhängt, ob sie A tat oder nicht.

Doch dieses inkompatibilistische Verständnis von Freiheit, so die kompatibilistischen Kritiker, ist inkonsistent: Einerseits soll die Handlung durch die vorangegangenen Geschehnisse nicht festgelegt sein; dieselbe Person mit denselben Wünschen, Charaktermerkmalen und Einstellungen hätte der inkompatibilistischen Auffassung zufolge unter exakt denselben äußeren Umständen auch anders handeln können. Andererseits soll es aber doch von dieser Person abhängen, was sie tut. Aber nichts an oder in dieser Person kann den Ausschlag dafür geben, A zu tun oder zu unterlassen, denn für jeden angeblich ausschlaggebenden Grund G soll ja gelten, dass die Person sich trotz G auch anders hätte entscheiden können. Doch wenn nichts an oder in der handelnden Person verständlich machen oder erklären kann, was die Person tut, dann hängt es auch nicht von dieser Person ab, was sie tut. Genau in dem Maße, in dem die Handlung durch vorhergehende Umstände, einschließlich der Einstellungen der handelnden Person selbst, nicht determiniert ist, ist sie der Kontrolle der Person entzogen. Nicht der Determinismus, so die Position des so genannten „harten“

Kompatibilismus, sondern der Indeterminismus ist mit Freiheit unvereinbar. Was die Inkompatibilisten als Freiheit bezeichnen, ist aus kompatibilistischer Sicht nichts anderes als Zufall. Umgekehrt ist die kompatibilistische Freiheit aus Sicht der Inkompatibilisten nur ein schlechtes Surrogat, denn was für eine Freiheit sollte das sein, die damit vereinbar ist, dass man keine Alternative hatte?

Im Laufe einer zweitausendjährigen Debatte haben die Vertreter beider Seiten ihre eigenen Positionen immer weiter verfeinert und weiterentwickelt und immer subtilere Argumente gegen die Auffassung der jeweils anderen Seite erdacht, doch bis heute hat sich keine Seite durchsetzen können. Auch die neueste, seit über zwanzig Jahren intensiv und auf hohem Niveau geführte Diskussion zwischen Kompatibilisten und Inkompatibilisten hat keiner Seite entscheidende Vorteile gebracht. Immer noch haben beide Positionen zahlreiche Anhänger, die sich ehrlicherweise zugestehen müssen, dass die eigenen Argumente offenbar nicht ausreichen, um die Anhänger der Gegenposition zu überzeugen.

Manche Autoren haben daraus den Schluss gezogen, dass der Begriff der Willensfreiheit inkonsistent ist.² Die Inkompatibilisten haben demnach darin Recht, dass Freiheit und Determinismus unvereinbar sind; die Kompatibilisten darin, dass Freiheit und Indeterminismus unvereinbar sind. Doch wenn Willensfreiheit weder mit dem Determinismus noch mit dem Indeterminismus vereinbar ist, dann muss der Begriff der Willensfreiheit in sich widersprüchlich sein.

Obwohl man zugeben muss, dass die Argumente beider Seiten gegen die jeweils andere Position erhebliche Kraft haben, scheint mir dies eine konsistente und sachliche angemessene Explikation des Begriffs der Willensfreiheit nicht auszuschließen. So könnten sich Freiheit und Determinismus insofern als vereinbar erweisen, als eine generelle Determinismusthese für die Freiheit des Willens einfach unerheblich ist. Diesem so genannten „weichen“ Kompatibilismus zufolge ist Freiheit sowohl mit dem Determinismus als auch mit dem Indeterminismus vereinbar. Auf die Grundzüge einer solchen Position werde ich gegen Ende dieses Beitrags zurückkommen. Es ist jedoch anzunehmen, dass auch diese Version des Kompatibilismus eingefleischte Inkompatibilisten nicht von ihrer Position abbringen wird. Der Grund dafür ist einfach: Auch die raffinierteste Version des Kompatibilismus muss zulassen, dass eine Handlung auch dann frei sein kann, wenn sie letztlich durch Umstände determiniert ist, die der Handlende selbst nicht kontrollieren kann, und genau das findet ein Inkompatibilist unannehmbar. Darin, so glaube ich, liegt der intuitive Kern des Inkompatibilismus: Wenn der Determinismus wahr sein sollte, dann sind alle unsere

² Vgl. z.B. Strawson (1986); dazu unten Abschnitt 3.

Entscheidungen vollständig durch Faktoren festgelegt, auf die wir keinen Einfluss haben. Doch Entscheidungen, die durch Faktoren festgelegt sind, auf die wir keinen Einfluss haben, sind nicht frei. Es ist die intuitive Plausibilität dieses Prinzips, die erklärt, warum der Common Sense nach Meinung vieler Philosophen auf Seiten des Inkompatibilismus steht. Im Folgenden möchte ich der Frage nachgehen, worauf diese Plausibilität beruht.

Meine Antwort wird lauten, dass die Plausibilität des Inkompatibilismus von einer bestimmten, keineswegs alternativlosen Konzeption von Vernunft abhängt. Wenn es stimmt, dass die inkompatibilistische Intuition auf einem anspruchsvollen Vernunftkonzept beruht, dann ist diese Intuition nicht theoretisch unschuldig und alternativlos. Indem wir die Plausibilität des Inkompatibilismus auf eine bestimmte Konzeption von Vernunft zurückführen, die sich ihrerseits als kritikwürdig erweist, kann die Verteidigung des Kompatibilismus über jenen toten Punkt hinauskommen, an dem der Inkompatibilist sich auf die scheinbar unabweisbare Intuition beruft, dass Handlungen nicht frei sein können, wenn sie durch für uns unkontrollierbare Faktoren determiniert sind.

Um das zu zeigen, werde ich im nächsten Abschnitt zwei klassische Argumente gegen die Vereinbarkeit von Freiheit und Determinismus vorstellen, nämlich das von Galen Strawson so genannte *Basic Argument* sowie das *Konsequenzargument* Peter van Inwagens, und zeigen, dass beide auf Schlussprinzipien beruhen, die sich bei näherem Hinsehen als strukturell identisch und inhaltlich eng verwandt erweisen.

3. Zwei Argumente für den Inkompatibilismus

Galen Strawson hat in seinem Buch *Freedom and Belief* sowie seither in verschiedenen Aufsätzen die These vertreten, dass Willensfreiheit und moralische Verantwortung unmöglich sind, und zwar unabhängig davon, ob der Determinismus wahr ist oder nicht.³ Dabei geht es Strawson um Verantwortung in einem starken Sinn, nämlich als Voraussetzung dafür, dass Menschen Strafe oder Belohnung wirklich *verdienen*. Um zu zeigen, dass es Verantwortung in diesem Sinn nicht gibt, hat Strawson eine Vielzahl von Argumenten vorgestellt, die seiner Meinung nach jedoch alle Varianten eines klassischen Grundarguments sind, des *Basic Argument*. Für unsere Zwecke reicht es aus, die intuitive Grundversion dieses Arguments zu betrachten, die Strawson folgendermaßen formuliert:⁴

Niemand ist Ursache seiner selbst (*causa sui*).

³ Vgl. u.a. Strawson (1986), (2002), (2003).

⁴ Vgl. Strawson (2003), 212.

Man ist nur dann für sein Handeln verantwortlich, wenn man zumindest in einigen wichtigen mentalen Hinsichten Ursache seiner selbst ist.

Niemand ist für sein Handeln verantwortlich.

Dieses Argument betrifft nicht unmittelbar die Freiheit des Willens, sondern die Möglichkeit moralischer Verantwortung. Doch Strawson macht klar, dass ein paralleles Argument gegen die Möglichkeit von Willensfreiheit vorgebracht werden kann. Das Argument ist formal schlüssig. Gesteht man die Wahrheit der Prämissen zu, dann folgt, dass Freiheit und Verantwortung unmöglich sind, und zwar unabhängig davon, ob der Determinismus wahr ist oder nicht.

Betrachten wir die Prämissen etwas näher. Die erste Prämisse besagt, dass niemand Ursache seiner selbst ist. Hier könnte ein Inkompatibilist Einspruch erheben. Natürlich hat jeder Mensch Eltern und ist insofern nicht Ursache seiner selbst. Aber gerade mit Blick auf diejenigen „wichtigen mentalen Hinsichten“, die in der zweiten Prämisse erwähnt werden, könnte es sehr wohl sein, dass wir uns in gewisser Weise selbst hervorbringen, ohne darin durch äußere Ursachen determiniert zu sein. Die mentalen Eigenschaften, auf die es in der zweiten Prämisse ankommt, sind der Charakter, die Wünsche, Werteinstellungen und Überzeugungen einer Person. Fassen wir alle diese mentalen Eigenschaften unter dem Begriff der *Persönlichkeit* zusammen, so könnte ein Inkompatibilist die These vertreten, dass wir unsere Persönlichkeit selbst hervorbringen, indem wir uns in bestimmten Situationen frei für oder gegen eine bestimmte Handlung entscheiden. Wenn ich mich entscheide, mit dem Rauchen aufzuhören, dann hat dieser Entschluss Auswirkungen auf meine zukünftigen Wünsche und Bedürfnisse und damit auf die weitere Entwicklung meiner Persönlichkeit; zugleich ist diese Entscheidung, sofern sie in einem inkompatibilistischen Sinn frei ist, nicht durch vorhergehende Faktoren verursacht. Insofern könnten wir also doch Ursache unserer selbst sein, und zwar in genau derjenigen Hinsicht, auf die es für die zweite Prämisse ankommt.

Strawson diskutiert diesen Einwand nicht explizit, weil er die These, dass niemand Ursache seiner selbst ist, für evident hält. Aus seinen Überlegungen geht jedoch hervor, dass er ihn folgendermaßen zurückweisen würde: Soll die freie Entscheidung, mit dem Rauchen aufzuhören, eine Entscheidung sein, für die ich verantwortlich bin, dann darf sie nicht rein zufällig zustande gekommen sein, sondern muss sich auf eine nachvollziehbare Weise aus meinen Wünschen, Einstellungen, kurz: aus meiner Persönlichkeit zum Zeitpunkt der Entscheidung ergeben. Doch damit verschiebt sich das Problem nur zurück in die Vergangenheit: Um für meine Persönlichkeit zum Zeitpunkt der Entscheidung verantwortlich

zu sein, müsste sie auf frühere freie Entscheidungen zurückgehen, die jedoch ebenfalls nur dann frei sein können, wenn sie Ausdruck einer Persönlichkeit sind, die auf noch frühere freie Entscheidungen zurück geht usw. Letztlich, so Strawson, gelangt man auf diese Weise immer zu Faktoren wie dem eigenen Erbgut und der frühkindlichen Erziehung, die man sicherlich nicht selbst hervorgebracht hat. Und da diese Faktoren in Verbindung mit den Erfahrungen, die man im Laufe seines Lebens macht, die eigene Persönlichkeit festlegen, ist man auch in dieser Hinsicht nicht Ursache seiner selbst.⁵

Damit kommen wir zur zweiten Prämisse, die besagt, dass man nur dann für sein Handeln verantwortlich ist, wenn man zumindest in einigen wichtigen mentalen Hinsichten Ursache seiner selbst ist. Wir können dies nun folgendermaßen reformulieren: Man ist nur dann für sein Handeln verantwortlich, wenn man selbst die Ursache, oder zumindest Mitursache, seiner eigenen Persönlichkeit ist. Strawson begründet diese Prämisse nur mit dem Hinweis darauf, dass es schließlich von der eigenen Persönlichkeit abhängt, was man tut. Betrachten wir den entsprechenden Schritt in einer längeren Version des *Basic Argument*.⁶

- (1) Interested in free action, we are particularly interested in actions that are performed for a reason [...].
- (2) When one acts for a reason, what one does is a function of how one is, mentally speaking [...].
- (3) So if one is to be truly responsible for how one acts, one must be truly responsible for how one is, mentally speaking, at least in certain respects.
- (4) But to be truly responsible for how one is, mentally speaking, in certain respects, one must have brought it about that one is the way one is, mentally speaking, in certain respects.
[...]
- (9) [...] True self-determination is impossible because it requires the actual completion of an infinite series of choices of principles of choice.
- (10) So true moral responsibility is impossible because it requires true self-determination, as noted in (3).

⁵ Ich bin nicht sicher, ob ein Inkompatibilist sich an dieser Stelle geschlagen geben muss. Alles kommt darauf, ob man verständlich machen kann, dass eine indeterminierte Entscheidung, die a fortiori auch nicht durch meine Persönlichkeit zum Zeitpunkt der Entscheidung determiniert ist, dennoch meine freie Entscheidung sein kann, für die ich verantwortlich bin. Ich bin in dieser Hinsicht skeptisch, möchte diesen Punkt aber hier nicht weiterverfolgen, sondern Strawson seine erste Prämisse um des Argumentes willen zugestehen.

⁶ Strawson (2003), 213.

Es geht mir hier vor allem um den Übergang von (2) zu (3). Das Wort „so“ deutet an, dass Satz Nummer (3) aus Nummer (2) folgen oder zumindest durch ihn begründet werden soll. Doch dieser Übergang, für den Strawson nicht weiter argumentiert, ist keineswegs zwingend. Jedenfalls folgt (3) nicht ohne weiteres aus (2). Diese Folgerung ergibt sich erst dann, wenn wir das folgende Prinzip hinzuziehen: Man ist nur für das verantwortlich, für dessen Ursachen man verantwortlich ist. Nennen wir dies das *Prinzip des kausalen Verantwortungsregresses* oder kurz *Regressprinzip*. Nun kann man in der Tat schließen:

Man ist nur für das verantwortlich, für dessen Ursachen man verantwortlich ist.

Was man aus einem Grund tut, ist eine Wirkung („a function“) dessen, wie man in mentaler Hinsicht beschaffen ist.⁷

Also ist man für das eigene Tun nur verantwortlich, wenn man dafür verantwortlich ist, wie man in mentaler Hinsicht beschaffen ist.

Und da man demnach für die eigene mentale Beschaffenheit nur dann verantwortlich ist, wenn man auch für *deren* Ursachen verantwortlich ist usw., ist man für sein Handeln nur dann verantwortlich, wenn man sich in mentaler Hinsicht selbst verursacht hat. In Verbindung mit der ersten Prämisse des *Basic Argument*, dass niemand Ursache seiner selbst ist, folgt dann unmittelbar, dass niemand für sein Tun verantwortlich ist.

Das Regressprinzip klingt zunächst ganz harmlos: Wenn mein Handeln z.B. durch meine Entscheidung verursacht ist, dann muss ich für meine Entscheidung verantwortlich sein, um für mein Handeln verantwortlich zu sein. So weit, so gut. Doch seine volle Kraft entfaltet dieses Prinzip erst in der wiederholten Anwendung. Auch meine Entscheidung hat schließlich Ursachen; ich bin dem Regressprinzip zufolge aber nur dann für meine Handlung verantwortlich, wenn ich auch für die Ursachen meiner Entscheidung verantwortlich bin, und für deren Ursachen, und für deren Ursachen usw. Auf diese Weise ergibt sich entweder ein unendlicher Regress der Ursachen, so dass ich klarerweise nicht für mein Handeln verantwortlich bin, oder es gibt eine erste, unverursachte Ursache meines Handelns. Nur falls es eine solche erste Ursache meines Handelns gibt und ich für sie verantwortlich bin, bin ich dem genannten Prinzip zufolge auch für mein Handeln verantwortlich.

Ich möchte die Plausibilität dieser Argumentation an dieser Stelle nicht weiter untersuchen, sondern zunächst nur festhalten, dass sie auf dem Regressprinzip beruht.⁸ Dieses

⁷ Dass Strawson unter „a function of“ tatsächlich einen kausalen Zusammenhang versteht, wird u.a. daraus deutlich, dass er im nächsten Satz fortfährt: „But to be truly responsible for how one is [...] one must have brought it about that one is the way one is“ (Hervorh. M. W.).

Prinzip hat eine dezidiert anti-kompatibilistische Tendenz, denn wenn der Determinismus wahr ist, dann hat jede unserer Handlungen Ursachen, die so weit zurückreichen, dass wir klarerweise nicht für sie verantwortlich sind. Und daraus würde sich nach dem Regressprinzip unmittelbar ergeben, dass wir für keine unserer Handlungen verantwortlich sind. Man wird diesem Prinzip eine gewisse intuitive Plausibilität nicht absprechen können. Ob wir es akzeptieren sollten, ist eine Frage, auf die ich zurückkommen werde.

Doch zunächst zu Peter van Inwagens einflussreichem *Konsequenzargument* gegen die Vereinbarkeit von Freiheit und Determinismus. Eine informelle Version des Arguments lautet in den Worten van Inwagens folgendermaßen:

Wenn der Determinismus wahr ist, dann ergeben sich alle unsere Handlungen als Konsequenzen aus den Naturgesetzen und aus Ereignissen in der entfernten Vergangenheit. Aber es hängt nicht von uns ab, was vor unserer Geburt vor sich ging, und ebenso wenig hängt es von uns ab, was die Naturgesetze sind. Deshalb hängen [wenn der Determinismus wahr ist] auch deren Konsequenzen (einschließlich unserer gegenwärtigen Handlungen) nicht von uns ab.⁹

Die Grundidee des Arguments geht bis auf Boethius zurück.¹⁰ Auch Kant hat sich seiner bedient. Aus der „Nothwendigkeit im Causalverhältnisse“, so Kant,

folgt, daß eine jede Begebenheit [...] unter der Bedingung dessen, was in der vorhergehenden Zeit war, nothwendig sei. Da nun die vergangene Zeit nicht mehr in unserer Gewalt ist, so muß jede Handlung, die ich ausübe, durch bestimmte Gründe, *die nicht in meiner Gewalt sind*, nothwendig sein, d.i. ich bin zu dem Zeitpunkte, darin ich handle, niemals frei.¹¹

Kants Argument beruht auf derselben Grundidee wie das Konsequenzargument van Inwagens: Meine Handlungen sind nur dann frei, wenn ich kontrollieren kann, welche Handlungen ich vollziehe – wenn meine Handlungen also, wie Kant es nennt, „in meiner Gewalt sind“ oder, wie van Inwagen sagt, wenn es „bei mir liegt“ („it’s up to me“), was ich

⁸ Damit dieses Prinzip auch nur halbwegs plausibel erscheint, muss man es allerdings so lesen, dass Verantwortung für Ursachen nicht nur Alleinverantwortung bedeutet: Man ist für seine Handlungen nur dann verantwortlich, wenn man für deren Ursachen zumindest *mitverantwortlich* ist.

⁹ Van Inwagen (1983), 16.

¹⁰ Vgl. Jäger (2006).

¹¹ AA 5: 94.

tue. Wenn meine Handlungen durch irgendwelche vorhergehenden Umstände determiniert sind, dann habe ich die Kontrolle über meine Handlungen nur dann, wenn ich auch die Umstände kontrolliere, die sie determinieren. Doch jene Umstände reichen, wenn der Determinismus wahr ist, bis weit in die Zeit vor meiner Geburt zurück. Dann, so die von Kant und van Inwagen gezogene Konsequenz, ist keine unserer Handlungen und Entscheidungen frei, weil wir keine von ihnen in unserer Gewalt haben.

Das besondere Verdienst van Inwagens liegt also nicht darin, ein völlig neues Argument gegen den Kompatibilismus entwickelt zu haben, sondern darin, dieses Argument in die Form einer deduktiv gültigen Ableitung gebracht zu haben, deren sechs Prämissen und zwei Schlussprinzipien offen zu Tage liegen¹² und auf ihre Plausibilität hin überprüft werden können. Tatsächlich haben Kritiker zeigen können, dass eines der beiden auf den ersten Blick hochplausiblen Schlussprinzipien, auf denen van Inwagens Argument beruht, das so genannte Prinzip β , ungültig ist.¹³ Alicia Finch und Ted Warfield haben jedoch eine neue und stark vereinfachte Version des Arguments entwickelt, die ohne das strittige Prinzip auskommt. In dieser Version lautet das Argument wie folgt:¹⁴

Niemand hatte jemals irgendeinen Einfluss darauf, welche Naturgesetze gelten und wie die Welt zu einem entfernten Zeitpunkt t_0 in der Vergangenheit beschaffen war.

Es gilt notwendigerweise, dass die Naturgesetze in Verbindung mit einer vollständigen Beschreibung der Welt zu t_0 implizieren, dass die Person P zu t_1 die Handlung H vollzieht.

Niemand hatte jemals irgendeinen Einfluss darauf, dass die Person P zu t_1 die Handlung H vollzieht.

Dieses Argument soll nachweisen, dass Freiheit und Determinismus unvereinbar sind. Die zweite Prämisse bringen die Inkompatibilisten daher nicht in eigenem Namen vor, sondern als Konsequenz aus dem Determinismus; in Verbindung mit der ersten Prämisse soll folgen, dass niemand jemals auf sein eigenes Handeln Einfluss hatte, was wiederum gleichbedeutend damit sei, dass niemand über Willensfreiheit verfügt.

Nehmen wir zunächst einmal an, dass es sich um ein schlüssiges Argument handelt und betrachten kurz die beiden Prämissen. Dass niemand, zumindest kein Mensch, jemals Einfluss auf die Naturgesetze hatte, scheint kaum bestreitbar zu sein. Dasselbe gilt für die

¹² Vgl. van Inwagen (1983), 93 ff.

¹³ Vgl. dazu Kapitan (2002), van Inwagen (2002), Jäger (2006).

¹⁴ Vgl. Finch/Warfield (1998), 521 f.

entfernte Vergangenheit. Wie steht es mit der zweiten Prämisse? Wenn wir unter dem Determinismus die These verstehen, dass die Naturgesetze in Verbindung mit einer vollständigen Beschreibung der Welt zu einem beliebigen Zeitpunkt notwendigerweise implizieren, wie die Welt zu jedem anderen Zeitpunkt beschaffen ist, dann folgt die zweite Prämisse aus der Annahme des Determinismus: Die Naturgesetze und der Zustand der Welt im Präkambrium legen dann eindeutig fest, welche Handlung ich gleich vollziehen werde.¹⁵ Und daraus scheint in Verbindung mit der ersten Prämisse zu folgen, dass ich keinen Einfluss darauf habe, welche Handlung ich gleich vollziehen werde.

Aber folgt dies wirklich? Schließlich handelt es sich hier nicht um einen einfachen Schluss nach dem *modus ponens*, denn die zweite Prämisse enthält einen Modaloperator, und die erste Prämisse sowie die Konklusion enthalten den Ausdruck „Niemand hatte jemals Einfluss darauf, dass...“, der in logischer Hinsicht ähnlich wie der Modaloperator „notwendigerweise“ funktioniert, aber nicht mit ihm identisch ist.¹⁶ Das Konsequenzargument in der hier betrachteten Form beruht auf folgendem Schlussprinzip:

¹⁵ Allerdings könnte man vielleicht bestreiten, dass die Naturgesetze mit Notwendigkeit gelten, was zur Folge hätte, dass die Naturgesetze und ein vergangener Weltzustand meine jetzige Handlung nicht mit Notwendigkeit implizieren würde; vgl. dazu Perry (2004).

¹⁶ Schlüsse, in deren Prämissen und Konklusionen Modaloperatoren vorkommen, können uns leicht in die Irre führen. Betrachten wir das folgende Beispiel:

Es gilt notwendigerweise, dass dann, wenn die Naturgesetze *NG* gelten und die Welt zu t_0 im Zustand Z_0 war, die Person *P* zu t_1 die Handlung *H* vollzieht.

Es gelten die Naturgesetze *NG* und die Welt war zu t_0 im Zustand Z_0 .

Es gilt notwendigerweise, dass die Person *P* zu t_1 die Handlung *H* vollzieht.

Dieser Schluss mag auf den ersten Blick gültig aussehen, er ist es aber nicht, denn er beruht auf dem Schlussprinzip: Notwendigerweise (wenn *p*, dann *q*); *p*; also notwendigerweise *q*. Dass dieses Prinzip nicht gilt, kann man sich an folgendem Beispiel klarmachen:

Es gilt notwendigerweise, dass dann, wenn England 1966 Fußballweltmeister wurde, Deutschland 1966 nicht Fußballweltmeister wurde.

England wurde 1966 Fußballweltmeister.

Es gilt notwendigerweise, dass Deutschland 1966 nicht Fußballweltmeister wurde.

Die Prämissen dieses Schlusses sind wahr, aber die Konklusion ist falsch: Es ist keineswegs notwendig, sondern bekanntlich bedingt durch eine höchst umstrittene Schiedsrichterentscheidung, dass Deutschland 1966 nicht Weltmeister wurde. – Ich will damit nicht behaupten, dass auch das Konsequenzargument auf einem einfachen Fehlschluss beruht, sondern nur anschaulich machen, wie wichtig es insbesondere bei Schlüssen mit

$Np, \Box(p \rightarrow q) \vdash Nq$

Daraus, dass niemand jemals Einfluss darauf hatte, dass p , und dass notwendigerweise gilt, dass wenn p , dann q , folgt, dass niemals jemand Einfluss darauf hatte, dass q .

Die intuitive Idee hinter diesem Prinzip lautet, dass die notwendigen Konsequenzen des Unbeeinflussbaren ebenfalls unbeeinflussbar sind. Wenn dieses Prinzip – nennen wir es in Anlehnung an John Martin Fischer das *Transferprinzip* – gültig ist, dann ist auch das Konsequenzargument gegen die Vereinbarkeit von Freiheit und Determinismus gültig.¹⁷

Mit Blick auf Strawsons *Basic Argument* hatte sich herausgestellt, dass es auf dem anti-kompatibilistischen Prinzip des kausalen Verantwortungsregresses beruhte: Man ist nur für das verantwortlich, für dessen Ursachen man verantwortlich ist. Die temporale Richtung des Regressprinzips geht von der Gegenwart in die Vergangenheit: Wenn ich für A verantwortlich bin, dann auch für die Ursachen von A. Durch die wiederholte Anwendung verfolgt das Regressprinzip die Bedingungen der Verantwortung immer weiter in die Vergangenheit zurück, bis wir schließlich bei einer Bedingung angelangt sind, die so weit zurückliegt, dass wir klarerweise nicht mehr für sie verantwortlich sein können.

Das Konsequenzargument hingegen verfährt genau umgekehrt: Es geht vom Zustand der Welt zu einem beliebigen Zeitpunkt in der fernen Vergangenheit aus, für den wir klarerweise nicht verantwortlich sind bzw. auf den wir keinen Einfluss hatten, und überträgt die Abwesenheit von Verantwortung bzw. Unbeeinflussbarkeit in die Zukunft. Dieser Transfer beruht auf dem Prinzip, dass die notwendigen Konsequenzen des Unbeeinflussbaren ebenfalls unbeeinflussbar sind. Doch obwohl sich beide Prinzipien in ihrer temporalen Richtung unterscheiden, beruhen sie auf derselben Grundstruktur. Um das zu sehen, beginnen wir mit dem Transferprinzip:

$(Np \ \& \ \Box(p \rightarrow q)) \rightarrow Nq$ ¹⁸

Dies ist logisch äquivalent mit der Aussage:

Modaloperatoren ist, genau darauf zu schauen, welches Schlussprinzip ihnen zugrunde liegt, und dessen Gültigkeit zu überprüfen.

¹⁷ Vgl. Fischer (1994).

¹⁸ Dies ist streng genommen nicht das oben genannte Transferprinzip T, welches eine Schlussregel ist, sondern dessen materiale Entsprechung. Für die Zwecke meiner Argumentation ist dieser Unterschied aber unerheblich.

$$(\text{non-N}q \ \& \ \Box (p \rightarrow q)) \rightarrow \text{non-N}p$$

Wenn es nicht der Fall ist, dass niemand jemals Einfluss darauf hatte, dass q , und notwendigerweise gilt, dass wenn p , dann q , dann ist es auch nicht der Fall, dass niemand jemals Einfluss darauf hatte, dass p .

Das kann man positiv auch so ausdrücken:

$$(\text{E}q \ \& \ \Box (p \rightarrow q)) \rightarrow \text{E}p$$

Wenn jemand Einfluss darauf hat, dass q , und notwendigerweise gilt, dass wenn p , dann q , dann hat auch jemand Einfluss darauf, dass p . M.a.W.: Nur auf diejenigen Ereignisse hat jemand Einfluss, auf deren Ursachen jemand Einfluss hat.

Um die logische Struktur des Regressprinzips deutlich zu machen, definieren wir zunächst den Begriff der Ursache folgendermaßen:

$$(U) \quad U(p, q) \stackrel{\text{Def.}}{=} \Box (p \rightarrow q) \text{ (für logisch voneinander unabhängige Aussagen } p \text{ und } q)$$

P ist genau dann die Ursache von q , wenn notwendigerweise gilt, dass p eine hinreichende Bedingung für q ist.¹⁹

Wir können das Regressprinzip nun so formulieren:

$$(\forall q \ \& \ \Box (p \rightarrow q)) \rightarrow \forall p$$

¹⁹ Dies ist eine stipulative Definition, die nicht beansprucht, die einzig richtige Analyse des Begriffs der Ursache auszurücken, sondern nur, eine von verschiedenen legitimen Verwendungsweisen des Ausdrucks „Ursache“ wiederzugeben. – Einer verbreiteten Auffassung zufolge ist Kausalität eine Relation ausschließlich zwischen Ereignissen, während andere Autoren auch Kausalrelationen zwischen Tatsachen zulassen. Ich möchte diese Frage hier offenlassen. Wenn ich im Folgenden der Einfachheit halber davon spreche, dass p Ursache von q ist, dann soll dies neutral gegenüber der gerade genannten Unterscheidung sein und zwei Lesarten zulassen: (1) die Tatsache, dass p , ist Ursache der Tatsache, dass q ; (2) die Tatsache, dass p , besteht darin, dass das Ereignis e_1 stattfindet; die Tatsache, dass q , besteht darin, dass das Ereignis e_2 stattfindet; und e_1 ist die Ursache (notwendigerweise hinreichende kausale Bedingung) von e_2 .

Wenn man für q verantwortlich ist, und p ist die Ursache für q , dann ist man auch für p verantwortlich. M.a.W.: Man ist für ein Ereignis nur dann verantwortlich, wenn man auch für seine Ursachen verantwortlich ist.

Betrachten wir nun noch einmal T3, das ja logisch äquivalent mit der ursprünglichen Formulierung des Transferprinzips T1 ist, dann stellen wir fest, dass es dieselbe logische Struktur hat wie das Regressprinzip:

$$(R) \quad (\forall q \ \& \ \Box (p \rightarrow q)) \rightarrow \forall p$$

$$(T3) \quad (E q \ \& \ \Box (p \rightarrow q)) \rightarrow E p$$

Beide Prinzipien übertragen eine bestimmte Eigenschaft (Verantwortung für bzw. Einfluss auf ein Ereignis) gleichsam rückwärts auf die Ursachen bzw. die notwendigerweise hinreichenden Bedingungen dieses Ereignisses: Man ist nur dann für q verantwortlich, wenn man für dessen Ursache p verantwortlich ist; es hat nur dann jemand Einfluss darauf, dass q , wenn auch jemand auf dessen Ursache p Einfluss hat.²⁰ Durch die wiederholte Anwendung dieser Prinzipien gelangen wir jeweils zu Ursachen unserer eigenen Handlungen, für die wir klarerweise nicht verantwortlich sind bzw. die wir nicht beeinflussen können, so dass folgt, dass wir auch für unsere eigenen Handlungen nicht verantwortlich sind bzw. sie nicht beeinflussen können. Diese Übereinstimmung ist insofern überraschend, als beide Prinzipien ja zunächst eine unterschiedliche „temporale Richtung“ haben und in ganz unterschiedlichen Argumenten mit unterschiedlichen Konklusionen eine Rolle spielen.

²⁰ Tatsächlich ist es nicht unplausibel, noch einen Schritt weiterzugehen und zu behaupten, dass beide Prinzipien äquivalent sind, sofern man zwei plausible Zusatzannahmen hinzuzieht:

Individualisierung von T3:

$$(T4) \quad (E_A q \ \& \ \Box (p \rightarrow q)) \rightarrow E_{AP}$$

Wenn Person A Einfluss darauf hat, dass q , und notwendigerweise gilt, dass wenn p , dann q , dann hat A auch Einfluss darauf, dass p .

Nun nehmen wir an:

$$(VE) \quad \forall p \ \Box E p$$

Man ist genau für das verantwortlich, worauf man Einfluss hat.

Nun können wir aus T4 die individualisierte Variante von R ableiten:

$$(R_i) \quad (\forall_A q \ \& \ \Box (p \rightarrow q)) \rightarrow \forall_{AP}$$

Wenn Person A für q verantwortlich ist, und p ist die Ursache für q , dann ist A auch für p verantwortlich.

Für meine Zwecke in diesem Beitrag reicht der Nachweis der Strukturgleichheit von R und T allerdings aus.

Wie ich nun im nächsten Abschnitt zeigen möchte, besteht eine zumindest auf den ersten Blick nicht weniger überraschende Übereinstimmung dieser beiden Prinzipien mit einem Prinzip, das Kant als das „oberste [...] Prinzip der reinen Vernunft“ bezeichnet. Wie ich argumentieren werde, zeigt sich darin die Abhängigkeit der inkompatibilistischen Schlussprinzipien von einer bestimmten, keineswegs alternativlosen Konzeption der Vernunft, deren inhärente Anfälligkeit für metaphysische Fehlschlüsse bereits Kant kritisiert.

4. Kant über den höchsten Grundsatz der reinen Vernunft

In der „Transzendentalen Dialektik“ der *Kritik der reinen Vernunft* geht es Kant darum, die Aussagen der traditionellen Metaphysik über die Seele, die Welt und Gott als Ergebnis von Fehlschlüssen zu erweisen, in welche sich die menschliche Vernunft notwendigerweise verwickelt, solange man die transzendentalphilosophische Unterscheidung zwischen Dingen an sich und Erscheinungen nicht angemessen berücksichtigt. Im Rahmen dieses Projekts, und zwar in der Auflösung der dritten von vier Vernunftantinomien, entwickelt Kant auch eine Lösung für das Problem der Vereinbarkeit von Freiheit und Determinismus, die Allen Wood treffenderweise als den Versuch bezeichnet hat, die Kompatibilität von Kompatibilismus und Inkompatibilismus nachzuweisen.²¹ Im Folgenden wird es jedoch nicht um Kants Auflösung der dritten Antinomie, sondern um seine Analyse der Grundstruktur vernünftigen Denkens gehen, die er seinem metaphysikkritischen Projekt in der „Einleitung zur transzendentalen Dialektik“ vorausschickt. Es ist diese Struktur, aus der sich Kant zufolge die metaphysischen Fehlschlüsse, darunter auch der auf die Unvereinbarkeit von Freiheit und Determinismus, ergeben.

Kant bestimmt die Vernunft im engeren, vom Verstand unterschiedenen Sinn, als „das Vermögen der Einheit der Verstandesregeln unter Prinzipien“²², d.h. die Fähigkeit, systematische Einheit in die Vielfalt unseres empirischen Wissens zu bringen.²³ Von dieser Fähigkeit können wir Kant zufolge entweder einen bloß logischen oder einen reinen bzw. transzendentalen Gebrauch machen. Im ersteren Fall ist die Vernunft lediglich die Fähigkeit, gegebenes Wissen zu ordnen, im zweiten Fall bringt sie selbst neues Wissen hervor – oder beansprucht dies zumindest. Kant will zeigen, dass der logische Gebrauch, sofern er nicht durch eine Kritik der reinen Vernunft im Zaum gehalten wird, unvermeidlich in den

²¹ Vgl. Wood (1984); zu Kants Auflösung der Freiheitsantinomie vgl. auch Willaschek (1992), Kap. III.

²² AA, 3: 239.

²³ Zum Folgenden vgl. ausführlicher Willaschek (2008).

transzendentalen Gebrauch umschlägt, der uns dann in die Fehlschlüsse der traditionellen Metaphysik führt.

Die systematische Einheit unseres empirisch bedingten Wissens, auf die der „logische Gebrauch der Vernunft“ abzielt, wird Kant zufolge hergestellt, indem verschiedene, zunächst unzusammenhängende Wissensinhalte auf dieselben Grundprinzipien zurückgeführt werden. Dies wiederum geschieht in Form von Vernunftschlüssen, die von einer gegebenen Aussage ausgehen und diese als Konklusion aus einem allgemeinen Obersatz und einem darunter fallenden Untersatz ableiten. In leichter Abwandlung eines Kantischen Beispiels lässt sich das folgendermaßen veranschaulichen: Wir wissen, dass alle festen Körper *veränderlich* sind. Außerdem wissen wir, dass alles *Zusammengesetzte* veränderlich ist. Wenn wir nun annehmen, dass alle Festkörper zusammengesetzt sind, dann haben wir eine Art Erklärung dafür, dass alle Körper veränderlich sind, nämlich, dass sie zusammengesetzt sind und alles Zusammengesetzte veränderlich ist. Und mehr als das: Wenn wir hinzunehmen, dass auch Gase zusammengesetzt und somit veränderlich sind, dann können wir feststellen, dass Festkörper aus *demselben* Grund veränderlich sind wie Gase, nämlich deshalb, weil sie zusammengesetzt sind. In diesem Sinn bringt die Vernunft durch logisches Schließen systematische Einheit in unser Wissen.

Diese Vernunftschlüsse steigen gleichsam aufwärts vom Bedingten zur Bedingung, in unserem Beispiel also von der Veränderlichkeit als Bedingtem zur Zusammensetzung als Bedingung. Entscheidend ist nun, dass dieser Aufstieg prinzipiell weiter fortgesetzt werden kann und die Vernunft auch darauf aus ist, ihn weiter fortzusetzen. In unserem Beispiel wäre etwa zum einen zu fragen, warum alles Zusammengesetzte veränderlich ist, zum anderen, warum alle Körper zusammengesetzt sind. Eine Kantische Antwort auf die letztere Frage könnte vielleicht lauten, dass alles Räumliche zusammengesetzt ist und alle Körper räumlich sind. Hier können wir nun weiterfragen, warum alles Räumliche zusammengesetzt ist, bis wir schließlich zu einem Obersatz kommen, der selbst nicht mehr aus einem noch allgemeineren Prinzip abgeleitet werden kann – zu einer Bedingung, die selbst nicht mehr bedingt ist: „der eigentümliche Grundsatz der Vernunft überhaupt (in ihrem logischen Gebrauche)“, so Kant, „sei: zu dem bedingten Erkenntnis des Verstandes das Unbedingte zu finden, womit die Einheit desselben vollendet wird“.²⁴ Dieses Prinzip ist lediglich eine „logische Maxime“, das heißt, eine Verfahrensregel vernünftigen Denkens, die für sich genommen keinerlei ontologische Implikationen hat. Ob es ein Unbedingtes wirklich gibt, in dem sich die Einheit der Vernunft vollendet, ist durch die bloße Anweisung, danach zu suchen, noch nicht

²⁴ AA, 3: 242.

ausgemacht. Dennoch gibt es Kant zufolge eine der menschlichen Vernunft selbst innewohnende Tendenz, von dieser vergleichsweise harmlosen logischen Maxime zu einem anspruchsvollen metaphysischen Prinzip überzugehen, das in der Annahme besteht: „wenn das Bedingte gegeben ist, so sei auch die ganze Reihe einander untergeordneter Bedingungen, die mithin selbst unbedingte ist, gegeben“²⁵, oder kurz: „Wenn das Bedingte gegeben ist, dann ist auch das Unbedingte gegeben.“ Dies, so Kant, ist „das oberste Prinzip der reinen Vernunft“²⁶, auf das alle transzendenten Sätze der klassischen Metaphysik zurückgehen.

Auf die zahlreichen exegetischen und sachlichen Fragen, die dieses Prinzip aufwirft, kann ich hier nicht eingehen. Mir geht es nur um jenen Aspekt der Kantischen Diagnose, der die logische Maxime ebenso betrifft wie den obersten Grundsatz, nämlich den *iterativen* Charakter von vernünftigen Begründungen und kausalen Erklärungen. Kants Rede von Bedingungen und Bedingtem in der Einleitung zur „Transzendentalen Dialektik“ erlaubt sowohl eine kausale als auch eine rationale Lesart. Anders gesagt: Kant legt sich nicht fest, ob mit „Bedingungen“ *Ursachen* oder *Gründe* gemeint sind; seine Analyse metaphysischer Fehlschlüsse soll beides umfassen. Und tatsächlich gilt ja für rationale Begründungen und kausale Erklärungen gleichermaßen, dass sie als Antworten auf Warum-Fragen ihrerseits weitere Warum-Fragen nach sich ziehen können. Insofern gilt in beiden Fällen, dass die Angabe einer, sei es kausalen, sei es rationalen, Bedingung die Frage nach der *Bedingung der Bedingung* erlaubt. Die von Kant herausgearbeitete „logische Maxime“ lässt sich somit folgendermaßen paraphrasieren: Wenn etwas Bedingtes gegeben ist, frage nach seiner Bedingung, und nach der Bedingung der Bedingung usw., bis du zu etwas Unbedingtem gelangst.

Die strukturelle Analogie zu Strawsons *Basic Argument* liegt auf der Hand. Nach Strawson bin ich für eine Handlung nur dann verantwortlich, wenn ich für deren Ursachen, und für die Ursachen der Ursachen usw. verantwortlich bin, was letztlich erfordert, dass es eine unverursachte Ursache meiner Handlung gibt, für welche ich verantwortlich bin. Noch deutlicher wird die Parallele, wenn wir die Konsequenzen betrachten, die Kants oberstes Prinzip der Vernunft für kausale Erklärungen bzw. rationale Begründungen hat. Der von Kant herausgearbeiteten Struktur rationalen Denkens zufolge ist ein Ereignis nämlich erst dann kausal vollständig erklärt, wenn auch seine Ursachen kausal erklärt sind, sowie deren Ursachen usw., bis wir zu einer unverursachten Ursache gelangen, die keine kausale Erklärung mehr erfordert und keine solche Erklärung erlaubt. Ganz analog ist eine Tatsache

²⁵ AA, 3: 243.

²⁶ Ebd.

demnach erst dann zur vollen Zufriedenheit unserer Vernunft begründet, wenn auch ihre Gründe rational begründet sind usw., bis wir zu einem Grund gelangen, der selbst nicht mehr begründet werden kann und nicht begründet werden muss. In Anlehnung an die obigen Formulierungen des Regress- und des Transferprinzips

$$(R) \quad (\forall q \ \& \ \Box (p \rightarrow q)) \rightarrow \forall p$$

$$(T3) \quad (Eq \ \& \ \Box (p \rightarrow q)) \rightarrow Ep$$

können wir diese Konsequenzen aus Kants „oberste[m] Prinzip der reinen Vernunft“ folgendermaßen formulieren:

$$(KE) \quad (KEq \ \& \ \Box (p \rightarrow q)) \rightarrow KEp$$

Falls p die Ursache von q ist, verfügen wir nur dann über eine vollständige kausale Erklärung für q , wenn wir auch über eine vollständige kausale Erklärung für p verfügen.

$$(RB) \quad (RBq \ \& \ \Box (p \rightarrow q)) \rightarrow RBp$$

Falls p der rational hinreichende Grund für q ist, verfügen wir nur dann über eine vollständige rationale Begründung für q , wenn wir auch über eine vollständige rationale Begründung für p verfügen.

Was diese Prinzipien miteinander gemeinsam haben, ist ihre *iterative Struktur*, d.h. ihre wiederholte Anwendbarkeit auf das Ergebnis früherer Anwendungen desselben Prinzips. Es ist das implizite „usw.“ dieser Prinzipien in Verbindung mit ihrem uneingeschränkten Geltungsanspruch für *alle* Paare von Ursachen und Wirkungen (Bedingungen und Bedingtem, Gründen und Begründetem), aus der sich das Dilemma von infinitem Regress oder unverursachter Ursache (unbedingter Bedingung, nicht mehr begründungsbedürftigem Grund) ergibt. Insofern sind sie alle diese Prinzipien Spezialfälle des „obersten Prinzips der reinen Vernunft“, zu jedem Bedingten eine entsprechende Bedingung anzunehmen.

Dass das *Basic Argument* Strawsons ohne den uneingeschränkt iterativen Charakter des Regressprinzips nicht in Gang kommen würde, kann man sich klarmachen, wenn man das Regressprinzip mit einem anderen, zunächst ganz ähnlich klingenden Prinzip vergleicht:

(APR) $(\forall q \ \& \ Kq \ \& \ \Box (p \rightarrow q)) \rightarrow \forall p$ (mit „Kq“ als „die Tatsache, dass q , besteht darin, dass die Person A die Körperbewegung K vollzieht“)

Wenn die Person A dafür verantwortlich ist, dass q , und die Tatsache, dass q , darin besteht, dass A eine Körperbewegung K vollzieht, und p die Ursache von q ist, dann ist A auch für p verantwortlich. M.a.W.: Man ist für seine Körperbewegungen nur dann verantwortlich, wenn man auch für deren Ursachen verantwortlich ist.

Der Unterschied zum Regressprinzip besteht darin, dass das alternative Prinzip eine wiederholte Anwendung nur dann erlaubt, wenn die Ursachen der Körperbewegungen ebenfalls Körperbewegungen sind, was nur sehr selten der Fall sein dürfte. Sofern die Ursache einer Körperbewegung zum Beispiel eine Entscheidung war, die ich getroffen habe, ergibt sich aus dem alternativen Prinzip nicht, dass ich für die Körperbewegung nur dann verantwortlich bin, wenn ich auch für die Ursachen meiner Entscheidung verantwortlich bin, da Entscheidungen keine Körperbewegungen sind.²⁷ Die Bedingungen für die Verantwortung für Entscheidungen könnten demnach ganz andere sein als die Bedingungen für Verantwortung für körperliche Handlungen (zum Beispiel, dass man die Entscheidung im Lichte von Gründen und Gegengründen getroffen hat, dass man nicht manipuliert oder auf andere Weise unzulässig beeinflusst worden ist usw.).

Dasselbe gilt für das Regressprinzip, auf dem das Konsequenzargument beruht. Betrachten wir das alternative Prinzip

(APT) $(\exists q \ \& \ Kq \ \& \ \Box (p \rightarrow q)) \rightarrow \exists p$

Wenn eine Person A Einfluss darauf hat, dass q der Fall ist, und die Tatsache, dass q , darin besteht, dass A eine Körperbewegung K vollzieht, und p die Ursache von q ist, dann hat A auch Einfluss darauf, dass p der Fall ist. M.a.W.: Man hat auf die eigenen Körperbewegungen nur dann Einfluss, wenn man auch auf deren Ursachen Einfluss hat.

APT taugt offensichtlich nicht als Schlussprinzip für das Konsequenzargument, denn anders als T1 bzw. T3 es erlaubt nicht den Schluss, dass dann, wenn ich weder die Vergangenheit noch die Naturgesetze beeinflussen kann, ich auch meine eigenen Körperbewegungen nicht

²⁷ Auch wenn man annimmt, dass Entscheidungen mit Vorgängen im Körper (z.B. mit Gehirnprozessen) identisch sind, handelt es sich dabei doch nicht um *Körperbewegungen* im gewöhnlichen Sinn.

beeinflussen kann. Dies folgt nach APT deshalb nicht, weil es nicht uneingeschränkt für *alle* Tatsachen gilt, dass ich sie nur dann beeinflussen kann, wenn ich ihre Ursachen beeinflussen kann, sondern nur für solche, die darin bestehen, dass ich eine Körperbewegung ausführe. Die Iteration, dass ich auch deren Ursachen beeinflussen können muss, und deren Ursachen usw. bis in die entfernte Vergangenheit, ergibt sich hier nicht, sofern es Ursachen meiner Körperbewegungen gibt, die nicht selbst Körperbewegungen sind (z.B. Entscheidungen). Für diese erfordert APT nicht, dass ich sie nur dann beeinflussen kann, wenn ich ihre Ursache beeinflussen kann, sondern lässt zu, dass die Bedingungen dafür, dass ich meine Entscheidungen beeinflussen kann (die Bedingungen dafür, dass meine Entscheidungen von mir abhängen), andere sind, als die Bedingungen dafür, dass ich meine Körperbewegungen (und über sie anderen Ereignisse in der Welt) beeinflussen kann.²⁸

Es geht mir an dieser Stelle nicht darum zu behaupten, dass APR und APT für sich genommen plausibler sind als R oder T, sondern nur um den Nachweis, dass es die unbegrenzte iterative Struktur der letztgenannten Prinzipien ist, aus der sich die inkompatibilistischen Konsequenzen ergeben. Beschränkt man die Iterierbarkeit auf einen begrenzten Bereich wie den der Körperbewegungen, so fallen auch die inkompatibilistischen Konsequenzen fort. Es ist die unbegrenzte Iteration, die uns, mit Kant gesprochen, vom Bedingten zur Bedingung und weiter zu deren Bedingung treibt und uns erst anzuhalten erlaubt, wenn wir bei etwas Unbedingtem angelangt sind. Und genau diese Dynamik, die den inkompatibilistischen Argumenten zugrunde liegt, charakterisiert Kant zufolge die menschliche Vernunft insgesamt und treibt sie in metaphysische Fehlschlüsse (unter anderem in die Freiheitsantinomie). Das Regress- und das Transferprinzip erweisen sich somit als

²⁸ APT soll nur belegen, dass es Alternativen zum Transferprinzip gibt, die keine uneingeschränkte Iteration erlauben. Eine inhaltliche Alternative zum Transferprinzip müsste jedoch auf einem plausiblen Begriff des Einflusshabens beruhen, der seinerseits nicht das Transferprinzip impliziert oder stützt. Das Folgende scheint mir ein aussichtsreicher Ausgangspunkt:

- (E) Eine Person A hat zum Zeitpunkt t_n genau dann Einfluss darauf, ob ein Ereignis E eintritt oder nicht, wenn es relativ zum Zustand der Welt zu t_n physikalisch möglich ist, dass eine Entscheidung von A zu t_{n+1} das Eintreten von E herbeiführt oder verhindert (bzw. schwächer: begünstigt oder erschwert).

Diese Formulierung impliziert nicht, dass man auf seine eigenen Handlungen nur dann Einfluss hat, wenn man auf alle ihre Ursachen und Bedingungen (einschließlich der Naturgesetze und der Vergangenheit) Einfluss hat, da die Frage, worauf man Einfluss hat, ausschließlich davon abhängt, welche Wirkungen *in der Zukunft* die eigenen Entscheidungen möglicherweise haben können.

Ausdruck derselben Konzeption von Vernunft und Begründung, die Kant in der Einleitung zur „Transzendentalen Dialektik“ analysiert. Weil diese Auffassung von Vernunft dadurch gekennzeichnet ist, dass ein Regress von Bedingtem und Bedingungen erst in einem *Unbedingten* zum Stillstand kommt, kann man von einer *absolutistischen* Konzeption von Vernunft sprechen. Sie ist dadurch gekennzeichnet, dass es nur zwei Weisen gibt, wie dem Anspruch der Vernunft genüge getan werden kann: durch etwas Unbedingtes oder aber durch etwas Bedingtes, das schrittweise auf etwas Unbedingtes zurückgeführt werden kann.

Wir sehen damit, dass die intuitive Plausibilität des Inkompatibilismus keineswegs theoretisch voraussetzungslos ist, sondern auf einer anspruchsvollen Konzeption von Vernunft, Begründung und Erklärung beruht. Im Rahmen dieses Beitrags ist es nicht möglich, diese absolutistische Vernunftkonzeption einer eingehenden Kritik zu unterziehen. Statt dessen möchte ich in aller Kürze zeigen, dass diese Konzeption nicht alternativlos ist und dass eine Alternative zu ihr mit einer Alternative zu den inkompatibilistischen Prinzipien R und T einhergeht. Wenn es eine solche Alternative gibt, dann führt das zu einer bedeutsamen Gewichtsverschiebung in der Debatte um den Kompatibilismus, da die scheinbar „intuitive“ Plausibilität des Inkompatibilismus demnach nicht nur auf einer bestimmten Konzeption von Vernunft beruht, sondern diese Konzeption sich zudem als gegenüber Alternativen begründungsbedürftig erweist.

5. Vernünftiges Fürwahrhalten aus praktischen Gründen

Wiederum ist mein Ausgangspunkt die Philosophie Kants, diesmal jedoch seine Postulatenlehre in der *Kritik der praktischen Vernunft*.²⁹ Dort will Kant die Rationalität des Glaubens an die Existenz Gottes und die Unsterblichkeit der Seele dadurch nachweisen, dass es sich um von ihm so genannte „Postulate der reinen praktischen Vernunft“ handelt. Unter einem solchen Postulat versteht Kant einen „*theoretischen*, als solchen aber nicht erweislichen Satz, sofern er einem a priori unbedingt geltenden *praktischen* Gesetze unzertrennlich anhängt“.³⁰ Es geht mir hier nicht um die Argumentation Kants im Einzelnen und auch nicht um die spezifischen Inhalte der Postulate, Gott und Unsterblichkeit, sondern allein um die Argumentationsfigur, derer sich Kant hier bedient. Ein Postulat der reinen praktischen Vernunft soll ein *theoretischer* Satz sein, d.h. ein Satz, der „sich auf den Gegenstand bezieh[t] und bestimm[t], was demselben zukomme oder nicht zukomme“.³¹ Als ein solcher

²⁹ Vgl. zum Folgenden ausführlicher Willaschek (im Erscheinen a), (im Erscheinen b).

³⁰ AA, 5: 122.

³¹ AA, 9: 110.

theoretischer Satz ist ein Postulat aber „unerweislich“, d.h. wir verfügen über keine theoretisch ausreichenden Gründe für ihn; allerdings auch nicht über ausreichende Gründe gegen ihn. Dieser Satz soll einem „a priori unbedingt geltenden *praktischen* Gesetze unzertrennlich anhängen“. Die Postulate sollen sich Kant zufolge nicht unmittelbar aus dem unbedingt geltenden Sittengesetz ergeben, sondern nur aus dem davon abgeleiteten Begriff des höchsten Gutes, welches in der proportionierten Verbindung von Moralität und Glückseligkeit besteht. Die Realisierung dieses höchsten Gutes ist ein notwendiges Ziel eines jeden vernünftigen Wesens. Doch seine Elemente Moralität und Glückseligkeit konfliktieren miteinander, weil, etwas vereinfacht gesagt, moralisch richtiges Handeln oft nicht glücklich macht und uneingeschränktes Glücksstreben zu unmoralischem Handeln führen kann. Nur wenn es Gott gibt und unsere Seelen unsterblich sind, so Kant, lässt sich dieser Konflikt auflösen, so dass dasjenige, was vernünftige Wesen notwendigerweise erstreben, zumindest prinzipiell erreichbar ist. In diesem Sinn also hängt ein Postulat der reinen praktischen Vernunft einem praktischen Gesetz „unzertrennlich an“: Es ist eine Bedingung dafür, dass wir uns einen widerspruchsfreien Begriff von uns selbst als vernünftig handelnden Wesen machen können, deren Handeln unbedingt geltenden praktischen Gesetzen untersteht.

Im Fall eines solchen Satzes, so Kant, ist es nun rational geboten, ihn als wahr anzunehmen, obwohl wir keine theoretisch ausreichende Begründung für ihn haben. Seine Begründung ist vielmehr praktischen Ursprungs, denn sie beruft sich auf die notwendigen Bedingungen unseres Selbstverständnisses als rationale Akteure. Zugleich ist seine Begründung aber insofern spezifisch vernünftig, als es um die Vermeidung von Widersprüchen in eben diesem Selbstverständnis geht. Es ist wichtig, dass diese Art von Begründung Kant zufolge einer rein theoretischen Gewissheit, was den Grad der Verlässlichkeit angeht, in nichts nachsteht: Die theoretische Vernunft darf und muss die Postulate annehmen „als ein ihr fremdes Angebot, das nicht auf ihrem Boden erwachsen, aber doch hinreichend beglaubigt ist“.³²

Treten wir einen Schritt zurück, so kann man die Grundidee der Kantischen Postulatenlehre folgendermaßen formulieren: Notwendige Implikationen unseres Selbstverständnisses als rational handelnde Wesen dürfen und müssen wir auch dann vernünftigerweise als wahr akzeptieren, wenn sie in rein theoretischer Hinsicht unbegründet sind. Auch wer Kants hyperrationalistische Moralbegründung nicht akzeptiert, wird zugeben müssen, dass es zahlreiche Überzeugungen gibt, die sich rein theoretisch nicht begründen lassen, die aufzugeben jedoch irrational wäre. Um einige eher unkantische Beispiele zu

³² AA, 5: 121.

nennen: Dass ein anderer Mensch mich liebt, dürfte niemals rein „theoretisch“ (d.h. aus einer neutral-detachierten Perspektive) feststellbar sein; und doch wäre es extrem unvernünftig, an der Liebe von mir nahe stehenden Menschen nur deshalb zu zweifeln, weil sie nicht theoretisch beweisbar ist. Ebenso verhält es sich mit der Möglichkeit sprachlicher Verständigung: Niemals ist ausgeschlossen, dass meine Worte missverstanden werden; doch kein vernünftiger Mensch wird allein aus diesem Grund aufhören zu glauben, dass ihn sein Gegenüber im Rahmen eines alltäglichen Gesprächs verstanden hat. Unzählige unserer Überzeugungen über uns und unsere Mitmenschen haben diesen Charakter: Sie sind rein theoretisch nicht begründbar, aber in praktischer Hinsicht unverzichtbar. Wir müssten an uns selbst als rationalen Wesen verzweifeln, wenn wir diese Überzeugungen aufgeben würden. Und genau deshalb wäre es irrational, sie aufzugeben (immer vorausgesetzt, dass keine überzeugenden theoretischen Gründe gegen sie sprechen).

Damit hat uns Kant selbst ein Paradigma vernünftiger Überzeugungsbildung an die Hand gegeben, das den Rahmen der Vernunftkonzeption der „Transzendentalen Dialektik“ sprengt. Es ist keineswegs so, dass nur solche Überzeugungen rational begründet sind, die selbst unbedingt gelten oder sich auf etwas Unbedingtes im Sinn der „Transzendentalen Dialektik“ zurückführen lassen. Zwar gilt auch das Sittengesetz, aus dessen Notwendigkeit sich die rationale Verbindlichkeit der Postulate ergibt, Kant zufolge „unbedingt“. Doch handelt es sich hier um die spezifisch *praktische* Geltung eines unbedingten (d.h. von kontingenten Neigungen und Zwecken unabhängigen) *Sollens*, nicht um eine *theoretische* Festlegung auf die Existenz ein unbedingten *Seienden* (Gott, das Weltganze, eine unsterbliche Seele). Kants Begründung der Postulate zeigt somit, dass vernünftige Überzeugungen *nicht* auf etwas im letzteren Sinn Unbedingtes zurückgehen müssen, denn Warum-Fragen können auch bei der Auskunft ein rational zufriedenstellendes Ende finden, dass etwas eine notwendige Bedingung vernünftigen Handelns ist. Die Bedingungen menschlichen Handelns erweisen sich somit als eine eigenständige Ressource rationaler Begründung.

Eine Auffassung von Vernunft, die solche nicht-absolutistischen Begründungsformen als nicht nur pragmatisch sinnvoll, sondern als für sich selbst gültig und rational hinreichend anerkennt, können wir als *kontextualistische* Konzeption von Vernunft bezeichnen, da dieser Auffassung zufolge dieselbe Überzeugung in *einem* (z.B. praktischen) Kontext ein nicht weiter begründungsbedürftiger Ausgangspunkt rationaler Begründungen oder kausaler Erklärungen sein kann, während man in einem *anderen* (z.B. theoretisch-spekulativen) Kontext durchaus nach einer Begründung oder Erklärung für sie fragen kann. Entscheidend ist dabei, dass die praktische Notwendigkeit einer Überzeugung auch dann ausreicht, sie

rationalerweise für wahr zu halten, wenn sie in rein theoretischer Hinsicht begründungsbedürftig, aber nicht weiter begründbar ist. Genau das besagt Kants These vom „Primat der reinen praktischen Vernunft“.³³ Allerdings beschränkt Kant, wie erwähnt, das praktisch Notwendige auf das, was dem „unbedingt geltenden“ Sittengesetz „unzertrennlich anhängt“, und die darauf gegründeten Überzeugungen zudem auf das, was selbst in einem ontologischen Sinn unbedingt ist, also Gott, Freiheit und die Unsterblichkeit der Seele. Insofern vertritt Kant selbst eine „kontextualistische“ Konzeption von Vernunft nur in einem sehr eingeschränkten Sinn, da er nur zwei mögliche Kontexte unterscheidet: den moralischen Kontext der reinen praktischen und den metaphysischen Kontext der „spekulativen“ Vernunft. Wenn wir Kants Position jedoch in der oben angedeuteten Weise verallgemeinern, gelangen wir zu einer Konzeption von Vernunft, die für unterschiedlichste Arten von Kontexten die Möglichkeit einräumt, dass etwas, das in einem Kontext ein unhinterfragbarer Ausgangspunkt ist, in einem anderen Kontext einer Begründung oder Erklärung bedarf.

6. Freiheit im Kontext

Es ist dieser kontextualistische Gedanke, der sich auch für die Vereinbarkeit von Freiheit und Determinismus fruchtbar machen lässt.³⁴ Die zentrale Idee lautet, dass sich der unendliche Regress zeitlich vorhergehender Ursachen unserer Handlungen dadurch vermeiden lässt, dass eine Ursache, nach deren Ursachen man in *einem* Kontext sinnvoll fragen kann, in einem *anderen* Kontext als *letzte* Ursache gelten kann, nach deren Ursachen nicht mehr sinnvoll gefragt werden kann. Ich schlage deshalb vor, zwischen zwei Arten von *Kontexten* und damit zwei Arten von Ursachen zu unterscheiden: *evaluative* Kontexte, in denen wir Personen für ihr Tun Verantwortung zuschreiben, und *explanative* Kontexte, in denen wir menschliches Verhalten wissenschaftlich erklären.

Betrachten wir zunächst die Bedingungen, unter denen wir einander im Alltag und im Recht für unser Tun verantwortlich machen, so stellen wir fest, dass diese Bedingungen in der Regel nicht positiv spezifiziert sind. Statt dessen gibt es eine Reihe von *Ausnahmebedingungen* – wie zum Beispiel mangelnde Reife, psychische Erkrankungen, Fremdsteuerung oder extreme emotionale Belastung –, unter denen Menschen für ihr Tun *nicht* verantwortlich sind. Umgekehrt gilt, dass man es dort, wo diese Bedingungen nicht vorliegen, mit einer Person zu tun hat, die für ihr Handeln (wenn auch nicht

³³ Vgl. 5: 119 ff.

³⁴ Vgl. hierzu ausführlicher Willaschek (im Erscheinen c).

notwendigerweise für alle seine Konsequenzen) verantwortlich ist. Dies wird besonders deutlich im deutschen Strafrecht, das Schuldfähigkeit ausschließlich negativ, über die Bedingungen für Schuldunfähigkeit (§ 20 StGB) und verminderte Schuldfähigkeit (§ 21 StGB) spezifiziert. Ganz im Einklang mit dem alltäglichen Verständnis von Verantwortung betreffen diese negativen Bedingungen in erster Linie Einschränkungen der Fähigkeit, sich im Lichte von Gründen und Gegengründen zu entscheiden und dieser Entscheidung gemäß zu handeln. Im Umkehrschluss kann man daraus folgern, dass für Verantwortung (und damit für Willensfreiheit als deren Voraussetzung) dem alltäglichen und strafrechtlichen Verständnis zufolge nicht mehr erforderlich ist als die *Fähigkeit*, sich nach Abwägung von Gründen (und insofern „vernünftig“) zu entscheiden und dementsprechend zu handeln.³⁵ Welche kausalen Bedingungen einen Menschen dazu in die Lage versetzen (etwa das Vorhandensein eines funktionsfähigen Gehirns, die Initiierung in eine bestimmte soziale Praxis usw.), spielt für die Zuschreibung von Verantwortung hingegen keine Rolle. Die einzigen Ursachen, die in *evaluativen Kontexten* (d.h. für die Frage, ob jemand für sein Tun verantwortlich ist) relevant sind, sind daher diejenigen, die Verantwortung *ausschließen*, indem sie vernünftige Entscheidungen oder ihre Umsetzung verhindern oder erschweren.

Wenn es hingegen um die naturwissenschaftliche Erklärung menschlichen Verhaltens geht, dann können offenbar auch ganz andere Ursachen relevant werden. Sicherlich werden die Einzelheiten des Stoffwechsels einer bestimmten Nervenzelle normalerweise nicht relevant sein, um das Verhalten eines Menschen zu erklären. Doch prinzipiell können in *explanativen Kontexten* alle kausalen Faktoren eine Rolle spielen. Auch die Frage nach den Ursachen der Ursachen, und nach deren Ursachen usw., ist hier prinzipiell sinnvoll und zulässig.

Sagen wir nun, dass eine Entscheidung frei ist, wenn sie selbst keine weiteren Ursachen hat, dann ist diese Formulierung zweideutig, solange wir nicht hinzufügen, um welchen Kontext es geht. Wenn wir annehmen, dass der Begriff der Willensfreiheit seinen logischen Ort nicht im Kontext naturwissenschaftlicher Erklärungen, sondern im praktischen Kontext der Zuschreibung von Verantwortung hat, dann muss es heißen: Eine Entscheidung ist frei, wenn sie keine Ursachen hat, *die in einem evaluativen Kontext relevant sind*. Wenn der Determinismus wahr sein sollte, dann gibt es keine absolut unverursachten Entscheidungen. Das muss die Möglichkeit freier Entscheidungen jedoch nicht ausschließen,

³⁵ Dabei bedeutet „vernünftig“ hier nicht, dass affektive und emotionale Faktoren keine Rolle spielen, sondern nur, dass diese in einen rationalen Abwägungsprozess einbezogen werden können, von dessen Ausgang die Entscheidung abhängt.

sofern eine Entscheidung als frei gilt, die keine Ursachen hat, die in einem evaluativen Kontext relevant sind – keine Ursachen, die Verantwortung dadurch ausschließen, dass sie vernünftiges Entscheiden und Handeln verhindern. Beispiele für solche in evaluativen Kontexten möglicherweise relevanten Ursachen sind Psychosen, Trunkenheit oder die Nachwirkungen frühkindlicher Vernachlässigung. Geistige Gesundheit, eine normale Erziehung, eine durchschnittliche Persönlichkeitsentwicklung und ein normal funktionierendes Gehirn, so sehr sie auch zu den Ursachen unseres Verhaltens gehören mögen, sind für die Frage der Verantwortung dagegen *insofern nicht relevant*, als ihr Vorliegen *Verantwortung weder sicherstellt noch ausschließt*. In einem evaluativen Kontext ist es demnach durchaus zutreffend, dass die allermeisten unserer Handlungen auf freie Entscheidungen zurückgehen, denn sie haben keine Ursachen, die in diesem Kontext (d.h. für die Zuschreibung von Verantwortung) *relevant* sind.

Es ist demnach die Unterscheidung zwischen relevanten und irrelevanten Ursachen, die den unendlichen Regress der natürlichen Ursachen beendet. Welche Ursachen relevant sind und welche nicht, ist keine ontologische Frage. Es ist vielmehr unsere menschliche, aber für jeden einzelnen Menschen unverfügbare Praxis, einander als verantwortliche Personen zu behandeln, die festlegt, welche Ursachen als Entschuldigungsgründe gelten und welche nicht. Und sobald wir über diese Unterscheidung verfügen, wird klar, dass unsere Entscheidungen normalerweise zwar Gründe, aber keine Ursachen haben, die in einem praktischen Kontext relevant sind. Freiheit und Determinismus sind demnach miteinander vereinbar, weil der Determinismus zwar impliziert, dass jedes Ereignis hinreichende Ursachen hat, nicht aber, dass jedes Ereignis hinreichende Ursachen hat, die in einem evaluativen Kontext relevant sind.

Es geht mir an dieser Stelle nicht darum, eine solche kontextualistische Konzeption von Freiheit plausibel zu machen und gegen mögliche Einwände zu verteidigen, sondern nur darum, die prinzipielle Verfügbarkeit einer Alternative zu einer inkompatibilistischen Freiheitskonzeption aufzuzeigen. Wie sich gezeigt hatte, beruhen zwei zentrale Argumente für den Inkompatibilismus auf Prinzipien, die genau jene Dynamik des Rückgangs vom Bedingten zur Bedingung in Anspruch nehmen, die Kant als Grundprinzip unserer Vernunft herausgearbeitet hat. Allerdings gibt es gute Gründe zu bezweifeln, dass es sich dabei wirklich um eine vollständige Charakterisierung menschlicher Rationalität handelt, da Kant selbst in der *Kritik der praktischen Vernunft* einen Typus rationaler Begründung entwickelt, der dem Dilemma von unendlichem Regress der Bedingungen und Annahme eines

Unbedingten entgeht, indem er das rationale Selbstverständnis handelnder Wesen, und damit das Praktische, als eigenständige Begründungsressource anerkennt.

So, wie die inkompatibilismusbegründenden Prinzipien R und T auf einer absolutistischen Konzeption von Vernunft beruhen, so geht die alternative kontextualistische Konzeption von Vernunft mit einer kompatibilistischen Auffassung von Freiheit einher, die es erlaubt, die Prinzipien R und T zurückzuweisen. Wenn eine Entscheidung dann frei und der Handelnde für sie dann verantwortlich ist, wenn sie keine Ursachen hat, die in einem evaluativen Kontext relevant sind, dann gilt weder, dass man nur für das verantwortlich ist, für dessen Ursachen man verantwortlich ist (R), noch, dass man nur das beeinflussen kann, dessen hinreichende Bedingungen man beeinflussen kann (T3). Tatsächlich erfordert die Verantwortung für die eigenen Entscheidungen der kontextualistischen Auffassung zufolge nur dann, dass man für die Ursachen der eigenen Entscheidungen verantwortlich ist, wenn diese Entscheidungen unter Ausschluss der eigenen rationalen Steuerungsfähigkeit zustande gekommen sind. Solche Fälle gibt es (etwa wenn man einen Zustand der Volltrunkenheit, der Verantwortung normalerweise ausschließen würde, selbst zu verantworten hat), aber sie sind selten. Es gilt demnach:

$$(NR) \quad (\Box (p \rightarrow q) \ \& \ EKRp) \leftrightarrow (Vq \ \Box \ Vp)$$

Dann und nur dann, wenn p eine in einem evaluativen Kontext relevante (d.h. Verantwortung ausschließende) Ursache von q ist, gilt, dass man für q nur verantwortlich ist, wenn man auch für p verantwortlich ist.

Dieses Prinzip, das die Falschheit von R impliziert, führt, anders als R, nicht in einen Regress der Ursachen, weil die allermeisten Ursachen dessen, wofür wir verantwortlich sind (insbesondere die Ursachen unserer eigenen Handlungen) in evaluativen Kontexten nicht relevant sind und wir daher für sie nicht verantwortlich sein müssen, um für unsere Handlungen verantwortlich zu sein. Normalerweise haben unsere Entscheidungen der kontextualistischen Auffassung zufolge keine Ursachen, für die man verantwortlich sein müsste, weil sie überhaupt keine Ursachen haben, die für die Frage der Verantwortung relevant sind. Aus demselben Grund gilt auch das Transferprinzip der kontextualistischen Auffassung zufolge nicht. Statt dessen gilt nur das eingeschränkte Prinzip:

$$(NT) \quad (\Box (p \rightarrow q) \ \& \ EKRp) \leftrightarrow (Eq \ \Box \ Ep)$$

Dann und nur dann, wenn p eine in einem evaluativen Kontext relevante (d.h. Verantwortung ausschließende) Ursache von q ist, gilt, dass man auf q nur dann Einfluss hat, wenn man auch auf p Einfluss hat.

Es liegt auf der Hand, dass NT nicht als Basis für das Konsequenzargument dienen kann, da unsere Handlungen und Entscheidungen nur in wenigen Ausnahmefällen Ursachen haben, die unsere rationale Steuerungsfähigkeit beeinträchtigen, somit Verantwortung ausschließen und in einem evaluativen Kontext relevant sind. Da insbesondere die Naturgesetze und der Zustand der Welt in der entfernten Vergangenheit nicht ausschließen, dass wir unsere Entscheidungen im Lichte von Gründen und Gegengründen treffen, ergibt sich nicht, dass wir unsere Entscheidungen nur dann beeinflussen können, wenn wir auch die Naturgesetze oder die entfernte Vergangenheit beeinflussen können.

7. Schluss

Die Konzeption von Vernunft, die Kant in der „Transzendentalen Dialektik“ analysiert und kritisiert hat, ist nicht ohne Alternative. War Kant in der *Kritik der reinen Vernunft* der Meinung, dass jede vernünftige Begründung entweder in einen unendlichen Regress mündet oder bei etwas Unbedingtem endet, macht er in der *Kritik der praktischen Vernunft* deutlich, dass Überzeugungen auch aufgrund ihrer praktischen Unverzichtbarkeit vernünftig sein können. Das „oberste Prinzip der reinen Vernunft“, wonach eine Erklärung nur dann vollständig ist, wenn sie auf eine unverursachte Ursache zurückgeführt werden kann, und eine Begründung nur dann abgeschlossen, wenn es ein oberstes, nicht weiter begründungsbedürftiges Prinzip gibt, gilt demnach nicht für die Vernunft als solche, sondern ist Ausdruck einer bestimmten, theoretisch keineswegs unschuldigen Konzeption von Vernunft. Eine Alternative zu dieser Auffassung besagt, dass Gründe nur dort erforderlich sind, wo praktisch relevante Fragen auftauchen, und dass Erklärungen nicht notwendigerweise bei einem Unbedingten enden, sondern bei dem, was in einem gegebenen Kontext unstrittig, notwendig oder offensichtlich ist. Ich behaupte nicht, dass Kant in der *Kritik der praktischen Vernunft* eine *solche* Konzeption von Vernunft vertreten hat. Aber seine Postulatenlehre bricht insofern mit der traditionellen Konzeption der Vernunft, als sie das Praktische als eigenständige Ressource rationaler Begründung anerkennt und so das Dilemma von Regress und Unbedingtem vermeidet.

Ich hatte zu zeigen versucht, dass zwei wichtige und, wie ich glaube, paradigmatische Argumente für den Inkompatibilismus auf strukturell identischen und inhaltlich eng

verwandten Prinzipien beruhen, die besagen, dass man nur für das verantwortlich ist, für dessen Ursachen man verantwortlich ist (R) bzw. nur das beeinflussen kann, dessen Ursachen man beeinflussen kann (T). Diese Prinzipien beruhen auf derselben Logik des Unbedingten, auf der auch die von Kant analysierte Konzeption von Vernunft beruht. So wie dieser Vernunftkonzeption zufolge nur das vollständig erklärt ist, dessen Ursachen erklärt sind, so ist man dem Inkompatibilismus zufolge nur für das verantwortlich, für dessen Ursachen man verantwortlich ist. Doch wenn diese Vernunftkonzeption nicht alternativlos ist, dann ist auch die damit verbundene Konzeption von Freiheit und Verantwortung nicht alternativlos.

Es war in diesem Beitrag nicht mein Ziel, eine solche alternative, d.h. kompatibilistische Konzeption von Freiheit und Verantwortung zu verteidigen. Hier wollte ich nur zeigen, dass es sowohl zur absolutistischen Konzeption von Vernunft als auch zu den inkompatibilistischen Prinzipien (R) und (T) konsistente und keineswegs unplausible Alternativen gibt. Die Intuitionen, auf denen der Inkompatibilismus beruht, sind somit genauso wenig theoretisch unschuldig wie die absolutistische Konzeption von Vernunft, sondern müssen gegenüber ihren Alternativen argumentativ ausgewiesen werden. Bereits die prinzipielle Verfügbarkeit einer alternativen Konzeption von Vernunft verschiebt die Gewichte in der Diskussion zwischen Kompatibilisten und Inkompatibilisten, denn die angeblich unmittelbar einleuchtenden Prinzipien, auf denen die inkompatibilistischen Argumente beruhen, erweisen sich nun als theorieabhängig und begründungsbedürftig.³⁶

³⁶ Ich danke Claudia Blöser, Sebastian Boll, André Fuhrmann, Thomas Höwing, Michael Quante und Peter Rohs für wertvolle Hinweise.

Siglen

Stellenangaben aus Kantischen Schriften beziehen sich mit Band- und Seitenzahl auf die Akademie-Ausgabe (Preußische Akademie der Wissenschaften (Hg.) (1900 ff.), *Kants Schriften*, Berlin).

Literaturverzeichnis

- Finch, Alicia und Warfield, Ted A. (1998), „The Mind Argument and Libertarianism“, in: *Mind* 107: 515-528
- Fischer, John Martin (1994), *The Metaphysics of Free Will: A Study of Control*, Oxford
- Geyer, Christian (Hg.) (2004), *Hirnforschung und Willensfreiheit*, Frankfurt
- Jäger, Christoph (2006), „Drei Konsequenzargumente für eine inkompatibilistische Theorie moralischer Verantwortung“, in: *Zeitschrift für Philosophische Forschung*, Band 60 (2006)
- Kane, Robert (Hg.) (2002), *The Oxford Handbook of Free Will*, Oxford
- Kapitan, Tomis (2002), „A Master Argument for Incompatibilism?“, in: Kane (2002)
- Keil, Geert (2007), *Willensfreiheit*, Berlin/New York
- Perry, John (2004), „Compatibilist Options“, in: Campbell, Joseph Keim, O'Rourke, Michael und Shier, David (Hg.) (2004), *Freedom and Determinism*, Cambridge
- Strawson, Galen (2003), „The Impossibility of Moral Responsibility“, in: Watson, Gary (Hg.) (2003), *Free Will*, Oxford
- Strawson, Galen (2002), „The Bounds of Freedom“, in: Kane (2002)
- Strawson, Galen (1986), *Freedom and Belief*, Oxford
- Van Inwagen, Peter (2002), „Free Will Remains a Mystery“, in: Kane (2002)
- Van Inwagen, Peter (1983), *An Essay on Free Will*, Oxford
- Willaschek, Marcus (im Erscheinen a), „The Primacy of Pure Practical Reason and the Very Idea of a Postulate“, in: Reath, A., Timmermann, J. (Hg.) (2009), *Kant's Critique of Practical Reason*, Cambridge
- Willaschek, Marcus (im Erscheinen b), „Rationale Postulate“, in Klemme, H. (Hg.) (2009), *Kant und die Zukunft der europäischen Aufklärung*, Berlin /New York
- Willaschek, Marcus (im Erscheinen c), „Non-Relativist Contextualism about Free Will“
- Willaschek, Marcus (2008), „Kant on the Necessity of Metaphysics“, in: *Proceedings of the 10th International Kant-Congress*, Berlin/ New York
- Willaschek, Marcus (1992), *Praktische Vernunft. Handlungstheorie und Moralbegründung bei Kant*, Stuttgart/Weimar
- Wood, Allen (1984), „Kant's Compatibilism“, in: Wood, Allen (Hg.) (1984), *Self and Nature in Kant's Philosophy*, Ithaca

Abstract

In diesem Beitrag gehe ich dem wechselseitigen Zusammenhang zwischen Vernunft und Freiheit nach, indem ich zu zeigen versuche, dass unterschiedliche Konzeptionen von Vernunft auch unterschiedliche Konzeptionen von Freiheit nach sich ziehen. Ich diskutiere zwei Argumente für den Inkompatibilismus (Strawsons *Basic Argument* und van Inwagens *Konsequenzargument*), von denen sich herausstellt, dass sie auf strukturgleichen Hintergrundannahmen beruhen, die ihrerseits Spezialfälle jenes Grundsatzes vernünftigen Denkens sind, den Kant als das „oberste Prinzip der reinen Vernunft“ bezeichnet. Diese („absolutistische“) Konzeption von Vernunft ist keineswegs alternativlos. Die „intuitive“ Plausibilität des Inkompatibilismus wird folglich in Frage gestellt.

In this paper I investigate the mutual link between reason and freedom by showing that different conceptions of reason entail different conceptions of freedom. I discuss two arguments for incompatibilism (Strawson's *Basic Argument* and van Inwagen's *Consequence Argument*) which turn out to rest on structurally identical background assumptions. These are instances of the basic principle which Kant calls the „highest principle of pure reason“. This (“absolutist”) conception of reason is not without alternative. Thus the “intuitive” plausibility of incompatibilism is called into question.