

SYMPOSIUM ZU: GEERT KEIL: WILLENSFREIHEIT*

Determinismus und Verantwortung: Was kann das Konsequenzargument?

Von CHRISTOPH JÄGER (Aberdeen/Innsbruck)

In seinem Buch *Willensfreiheit* hat Geert Keil jüngst eine libertarische Auffassung der Beziehung zwischen Willensfreiheit und Determinismus verteidigt und eine Reihe gängiger Argumente für kompatibilistische und deterministische Positionen kritisiert. Das Buch hat einen klaren Aufbau, liefert hilfreiche Zusammenfassungen zentraler Aspekte neuerer Debatten über Willensfreiheit und ist – darin liegt seine besondere Stärke – dezidiert argumentativ gehalten. Die in der von Keil skizzierten libertarischen Position in Anspruch genommenen inkompatibilistischen Intuitionen sind stark und halten meines Erachtens auch bei näherer philosophischer Analyse, was sie versprechen: Sie lassen sich zu guten Argumenten dafür ausbauen, dass Willensfreiheit und Determinismus unvereinbar sind.

Die Rekonstruktionen einiger wichtiger Stationen jüngerer Debatten über das Thema und einige zentrale systematische Überlegungen in dem Buch dürften allerdings zum Teil etwas zu flüchtig ausgefallen sein, um Kompatibilisten zu beeindrucken und der Stärke der inkompatibilistischen Position gerecht zu werden. Dies gilt zum Beispiel für Keils Diskussion des so genannten Konsequenzarguments für den Inkompatibilismus (83–87), das in den zeitgenössischen Diskussionen über Willensfreiheit zumindest im anglo-amerikanischen Raum seit fast drei Jahrzehnten eine herausragende Rolle spielt und bis heute von vielen Teilnehmern der Debatte als das Flaggship des Inkompatibilismus (und somit auch des Libertarismus) angesehen wird. „The most influential defenses of incompatibilism“, so betonen beispielsweise Thomas Crisp und Ted Warfield in einem wichtigen Aufsatz, „have employed some version of the Consequence argument“ (Crisp u. Warfield 2000, 173). Crisp und Warfield sind Inkompatibilisten, doch auch Kompatibilisten halten eine detaillierte Auseinandersetzung mit dem Konsequenzargument für unumgänglich. Tomis Kapitan etwa schreibt in seinem Artikel *A Master Argument for Incompatibilism?* über das Konsequenzargument: „The past twenty-five years have witnessed a vigorous discussion of an argument directed against the compatibilist approach to free will and responsibility“, und wenige Seiten später heißt es, der Weg, den Kompatibilisten letztlich beschreiten müssten, bestehe darin, das Konsequenzargument im Zuge einer geeigneten Theorie [*account*] der relevanten praktischen Modalitäten anzugreifen (Kapitan 2002, 127, 133). Wenn diese Einschätzung richtig ist, dann sind Inkompatibilisten (und somit auch Libertarier) entsprechend dazu angehalten, ihre Auffassung im Rahmen einer

* GEERT KEIL: WILLENSFREIHEIT. Walter de Gruyter, Berlin 2007, 222 S., ISBN 978-3-11-019561-3.

möglichst transparenten und gegen kompatibilistische Kritiken geharnischten Formulierung des Konsequenzarguments zu verteidigen.

Die einflussreichsten jüngeren Darstellungen und Verteidigungen des Konsequenzarguments stammen von Peter van Inwagen. Andere Formulierungen beziehungsweise Weiterentwicklungen oder kritische Diskussionen finden sich (außer bei Crisp und Warfield) zum Beispiel bei David Wiggins (1973), Carl Ginet (1966, 1990), David Widerker (1987), David Lewis (1993), Tim O'Connor (1993), Thomas McKay und David Johnson (1996), Robert Kane (1996), John Martin Fischer und Mark Ravizza (1998) und vielen anderen. Keil allerdings hält nicht viel vom Konsequenzargument, jedenfalls nicht in den Formen, wie es in zeitgenössischen Debatten diskutiert wird. Insbesondere kritisiert er den Rekurs von van Inwagens Konsequenzargument auf die Unbeeinflussbarkeit der fernen Vergangenheit: Dies sei ein „Umweg“ mit der Gefahr, „dass der Adressat [...] verwirrt wird, weil es ihm dadurch schwer gemacht wird, die spezifisch inkompatibilistischen Teile des Arguments zu identifizieren und von den unkontroversen zu unterscheiden“ (85). Van Inwagens Bezug auf die ferne Vergangenheit, so Keil, sei seines Erachtens eine „rhetorische Dramatisierung, die den Kompatibilisten mit ins Boot holen soll“. Denn „dass die Vergangenheit sich nicht ändern lässt, gehört ja nicht zu den Dingen, die zwischen Kompatibilisten und Inkompatibilisten umstritten sind“:

„Indem van Inwagen die Unbeeinflussbarkeit der Vergangenheit ins Spiel bringt, verbindet er den Determinismus mit einer Behauptung, die für diesen gar nicht spezifisch ist. Dieser Missgriff ist allerdings weit verbreitet und findet sich zum Beispiel auch bei Kant, welcher den Determinismus nur deshalb als freiheitsgefährdend ansieht, weil dann Handlungen ‚ihre bestimmende Gründe in der vorhergehenden Zeit haben (die mit dem, was sie in sich hält, nicht mehr in unserer Gewalt ist)‘.“ (Ebd.)¹

Nun kommen zweifellos auch unter Meistern einer Zunft wie Kant und van Inwagen gelegentlich Missgriffe vor. Aber in diesem Fall beruht der Vorwurf des Missgriffs auf einem Missverständnis des Kritikers. Im Folgenden werde ich in konstruktiver Absicht einige Missverständnisse und Unklarheiten in Keils Erörterung des Konsequenzarguments diskutieren, die dieses schwächer erscheinen lassen, als ihm gebührt. Meine Bemerkungen betreffen neben (i) Keils Kritik an dem Rekurs des Arguments auf die (ferne) Vergangenheit; (ii) seine Darstellung der allgemeinen Struktur und Dialektik des Konsequenzarguments; (iii) seine Verwendung des Begriffs der Notwendigkeit; sowie (iv) das allgemeine Ziel der inkompatibilistischen Überlegung, das letztlich in dem Aufweis der Unvereinbarkeit von Determinismus und *moralischer Verantwortlichkeit* liegt. Im ersten Abschnitt diskutiere ich (i), im zweiten Abschnitt Punkt (ii), im dritten Abschnitt einen Alternativvorschlag Keils sowie nochmals (i) und (ii), und im vierten Abschnitt betrachte ich einige Konsequenzen von Punkt (iv). Der vorliegende Rahmen gestattet es nicht, das Konsequenzargument umfassender zu verteidigen. Ziel meiner Überlegungen ist es an dieser Stelle vielmehr lediglich, seine Grundidee nebst einigen Stärken und Schwächen darzustellen. Der Tenor meiner Ausführungen lautet, dass man den Keil, den das Konsequenzargument zwischen Determinismus und Verantwortung setzt, um einiges tiefer treiben sollte, als es in dem Buch geschieht.

¹ Für Kants Argument in diesem Zusammenhang vgl. auch Willaschek (2008). Willaschek argumentiert, dass das Konsequenzargument beziehungsweise genauer: die Variante einer wichtigen Schlussregel, die es verwendet, strukturelle Parallelen zu Kants „oberstem Prinzip der reinen Vernunft“ aufweist, was das Argument in Willascheks Augen schwächt.

I. Die Rolle der fernen Vergangenheit

Das Konsequenzargument, so schreibt Keil unter Verweis auf Peter van Inwagen und Carl Ginet, bestehe aus zwei Prämissen und einer Konklusion und habe die folgende Form:

- „(P1) Wenn der Determinismus wahr ist, folgen unsere Handlungen aus Naturgesetzen und Ereignissen der fernen Vergangenheit.
- (P2) Es steht nicht in unserer Macht, die Naturgesetze zu ändern, noch die Ereignisse vor unserer Geburt.
- (K) Also stehen auch die kausalen Konsequenzen der Vergangenheit und der Naturgesetze nicht in unserer Macht, unsere eigenen Handlungen eingeschlossen.“ (Es folgt die Fußnote: „Vgl. van Inwagen 1983, 16 und 56; ähnlich Ginet 1966.“)

Die Rede von „folgen“ in (P1) ist etwas unglücklich, denn sie suggeriert ein logisches Verhältnis zwischen Handlungen und Naturgesetzen sowie Ereignissen der fernen Vergangenheit. Doch aus Ereignissen folgt nichts. Weiter unten werde ich Fassungen des Konsequenzarguments betrachten, welche die Überlegung nicht auf der ontologischen Ebene von Ereignissen, sondern im Rekurs auf Propositionen beziehungsweise Sätze, die bestimmte Propositionen ausdrücken, formulieren. Orientieren wir uns jedoch zunächst am Wortlaut von van Inwagens provisorischer Skizze seines Konsequenzarguments, auf die Keil sich beruft und in der Handlungen als (kausale) Konsequenzen von Naturgesetzen und Ereignissen der fernen Vergangenheit beschrieben werden:

„If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us.“ (van Inwagen 1983, 16 u. 56)

Wo liegt der angebliche „Missgriff“ dieser Überlegung, wenn sie Ereignisse in der fernen Vergangenheit ins Spiel bringt? Dem (Laplace-)Determinismus zufolge, so betont Keil, legt „jeder beliebige Weltzustand gemeinsam mit den Naturgesetzen jeden anderen Weltzustand [fest]. Laplace-Determiniertsein ist eine zeitsymmetrische Angelegenheit“ (84 f.). Verleitet durch die kausale Konnotation von „determinieren“, nehme man gewöhnlich an, dass die Vergangenheit die Gegenwart festlege; doch dem Determinismus zufolge könne „man im gleichen Sinn von ‚festlegen‘ sagen, dass die Gegenwart die Vergangenheit festlegt. Man darf also die Beziehung des Laplace-Determinierens nicht mit der Kausalitätsbeziehung verwechseln“ (85). Das ist zwar richtig; aber van Inwagen sind diese Sachverhalte nicht entgangen. Seine offizielle Charakterisierung von Determinismus lautet:

„We shall apply this term to the conjunction of these two theses: For every instant of time, there is a proposition that expresses the state of the world at that instant; [...] If p and q are any propositions that express the state of the world at some instants, then the conjunction of p with the laws of nature entails q .“ (van Inwagen 1983, 65)

Dies, so van Inwagen, erfasse wenigstens eine These, die man mit Recht „Determinismus“ nennen könne. Und explizit hebt er in diesem Zusammenhang hervor: „This definition has the consequence [...] that the future determines a unique past.“ (Ebd.) Wenige Sätze später betont er zudem: „The reader will note that the horrible little word ‚cause‘ does not appear in this definition.“ (Ebd.) Stellen wir vor diesem Hintergrund noch einmal Keils Frage: „Welche

Funktion hat also van Inwagens Bezug auf die ferne Vergangenheit bzw. auf Ereignisse vor meiner Geburt?“ (85)

Van Inwagen entwickelt (in van Inwagen 1983) drei detailliertere Fassungen seines Arguments. Ich werde hier aus Platzgründen nur die dritte, so genannte modale Fassung betrachten (ebd., 93–105). Diese will Folgendes zeigen:

- (T) Wenn der (kausal-naturgesetzliche) Determinismus wahr ist, dann gilt für eine beliebige wahre Proposition p : Niemand (kein menschliches Wesen) hat oder hatte jemals eine Wahl in Bezug die Wahrheit von p .

Dass ein Akteur S eine Wahl in Bezug auf eine wahre Proposition hat oder hatte, versteht van Inwagen in diesem Zusammenhang so, dass S dafür sorgen oder sicherstellen kann beziehungsweise konnte, dass nicht- p . Unten diskutiere ich eine präzisiertere Fassung des Arguments für (T). Zunächst jedoch gilt es festzuhalten, dass van Inwagens Argument für diese These (kontra Keil: 85) nicht einfach die Unbeeinflussbarkeit der Vergangenheit ins Spiel bringt, sondern sich darauf beruft, dass niemand eine Wahl in Bezug auf Weltzustände hat *oder je hatte*, die lange vor der Existenz menschlicher Wesen herrschten. Der Sinn dieses Vorgehens liegt schlicht darin, dass das Argument auf diese Weise Komplikationen vermeidet, die für seine Grundidee nicht von Belang sind. Betrachten wir zunächst, welche Konsequenzen es für das Konsequenzargument hätte, wenn van Inwagen, statt sich auf die ferne Vergangenheit zu berufen, die unspezifischere Formulierung gewählt hätte: „Wenn der Determinismus wahr ist, sind unsere Handlungen die Konsequenzen von jeweils vergangenen Ereignissen und Naturgesetzen.“ Diese Behauptung wäre zwar nicht falsch; aber sie wäre für van Inwagens inkompatibilistisches Argument ungeeignet, weil sie nicht ausschliesse, dass unter den betreffenden *vergangenen* Ereignissen solche sind, in Bezug auf die wir sehr wohl eine Wahl hatten. Und in diesem Fall könnte nicht auf (T) geschlossen werden. Denn diejenigen Handlungen eines Akteurs S , die Konsequenzen von Naturgesetzen und solchen Ereignissen sind, in Bezug auf *die* S einmal eine Wahl hatte, lassen sich nicht plausibel als Handlungen betrachten, in Bezug auf die S weder eine Wahl hat *noch jemals hatte*. Betrachten wir dazu das folgende Beispiel van Inwagens:

„Suppose that when I am drunk it is not within my power to refrain from violently assaulting those who disagree with me about politics. I get drunk and overhear a remark about Cuban troops in Angola and, soon, therefore, Fred’s nose is broken. I was, under the circumstances, unable to refrain from breaking Fred’s nose. And yet no one doubts that I am to blame for his broken nose. How can that be? Simple: Although I was unable to avoid breaking his nose, that inability is one I could have avoided having.“ (van Inwagen 1989, 236)

Angenommen, van Inwagen betrank sich zu t_1 , hätte dies jedoch auch unterlassen können. Angenommen, seine Attacke gegen Fred erfolgt zu t_2 . Laut Hypothese kann van Inwagen dann zwar zu t_2 nicht anders, als Fred anzugreifen; er hat in Bezug hierauf keine andere Wahl. Doch daraus folgt nicht, dass er zu t_2 weder eine Wahl in Bezug auf sein (dann gegenwärtiges) Verhalten hat *noch jemals hatte*. Denn van Inwagen hatte eine Wahl hinsichtlich dessen, was er zu t_1 tat und was ihn dazu determinierte, Fred zu t_2 zu attackieren. Der betreffende Punkt lässt sich auch wie folgt beschreiben. Das Prinzip:

Wenn ein gegebener Weltzustand vergangen ist, dann hat oder hatte niemand jemals eine Wahl in Bezug auf diesen Zustand

ist falsch. Denn aus der bloßen Tatsache, dass ein Weltzustand vergangen ist, folgt nicht, dass niemand *jemals* dafür sorgen konnte, dass er nicht eintreten würde.

Analoges gilt auch für Ereignisse beziehungsweise Zustände, die nicht vor, sondern während der zu bewertenden Handlungen oder später als diese stattfinden oder vorliegen. Angenommen, das Konsequenzargument behauptete, statt sich auf die (ferne) Vergangenheit zu berufen: „Wenn der Determinismus wahr ist, sind all unsere Handlungen durch Naturgesetze und von diesen Handlungen aus gesehen zukünftige Ereignisse festgelegt.“ Auch mit dieser Prämisse ließe sich die gewünschte inkompatibilistische Konklusion nicht ableiten, denn auch der Verweis auf von der Handlung eines Akteurs *S* aus gesehen zukünftige Ereignisse schlosse ohne weitere Qualifikationen nicht aus, dass *S* Einfluss auf diese Ereignisse und eine Wahl in Bezug auf ihr Eintreten hatte (hat oder einmal haben wird), womit ebenfalls nicht ohne weiteres gefolgert werden könnte, dass *S* hinsichtlich beliebiger Handlungen keine andere Wahl hatte, als sie eintreten zu lassen. Bezieht man sich in dem Argument hingegen auf einen beliebigen vollständigen Weltzustand in grauer Vorzeit, dann bedarf die betreffende Prämisse keiner Einschränkungen, die die Unabhängigkeit von menschlichen Handlungen sicherstellen. Dass kein menschliches Wesen jemals eine Wahl in Bezug auf Ereignisse hatte oder je haben wird, die im Präkambrium stattfanden, bedarf keiner weiteren Erläuterungen oder zusätzlicher Argumente.

Obwohl somit einleuchtet, warum van Inwagen – mit guten Gründen – die ferne Vergangenheit ins Spiel bringt, trifft Keils Unbehagen dennoch einen interessanten Punkt. In dem Bezug des Konsequenzarguments auf die ferne Vergangenheit liegt womöglich eine Schwäche, allerdings aus anderen als den von Keil vermuteten Gründen. Das Konsequenzargument tritt als ein allgemeines Argument für den Inkompatibilismus an. Wenn man diesen jedoch als die These versteht, dass Determinismus und Willens- beziehungsweise Wahlfreiheit *notwendigerweise*, in jeder möglichen Welt, unvereinbar sind², dann lässt sich mit einer Überlegung, die in einer ihrer Prämissen die Existenz einer entfernten Vergangenheit voraussetzt, nicht für die Wahrheit des Inkompatibilismus argumentieren. Dass es in der aktuellen Welt eine entfernte Vergangenheit gibt, ist eine kontingente Eigenschaft dieser Welt. Anderen Welten, unter ihnen auch solchen mit handelnden Personen, fehlt dieses Charakteristikum. Was somit van Inwagens modales Konsequenzargument, falls stichhaltig, genau genommen lediglich zeigt, ist, dass in einer Welt, die sich zeitlich in eine entfernte Vergangenheit erstreckt, der Determinismus mit der Existenz von Willensfreiheit unvereinbar ist.³ Die allgemeinere These, der zufolge es keine mögliche Welt gibt, in der Willensfreiheit und Determinismus gemeinsam existieren, beweist es nicht.

Dieser Punkt ist wichtig zum Verständnis des Konsequenzarguments, doch ist er meines Erachtens nicht desaströs für den Libertarier. Das Problem entsteht, *wenn* man den Inkompatibilismus als die These versteht, dass Willensfreiheit und Determinismus notwendigerweise unvereinbar sind. Was den Freiheitstheoretiker jedoch in erster Linie motiviert, ist die Frage, ob Freiheit und Verantwortung in *unserer*, der aktuellen Welt und für die in ihr handelnden Personen auch dann existieren, wenn Handlungen und Entscheidungen in dieser Welt kausalnaturgesetzlich determiniert sind. Nennen wir die These, dass Freiheit und Determinismus sich notwendigerweise ausschließen, *starken Inkompatibilismus*; die schwächere Behauptung dagegen, dass, wenn der Determinismus wahr ist, (*de facto*) keine Freiheit existiert, sei als *moderater Inkompatibilismus* bezeichnet. Da es (trotz der Möglichkeit skeptischer Zweifel) rational ist anzunehmen, dass die aktuelle Welt eine Welt mit entfernter Vergangenheit ist,

² Dass eigentlich dies die korrekte inkompatibilistische These sei – entgegen dem, was viele Verfechter des Konsequenzarguments zu zeigen versuchen –, meint beispielsweise Warfield (2000).

³ Hierauf hat neuerlich Campbell (2007) hingewiesen.

kann van Inwagens modales Konsequenzargument, sofern es auch in seinen anderen Teilen akzeptabel ist, nach wie vor als ein gutes Argument für einen moderaten Inkompatibilismus angesehen werden. Diese Konklusion, obwohl metaphysisch schwächer als die These, dass Determinismus und Freiheit unmöglich zusammen existieren können, ist interessant genug.

II. Zwei Konsequenzargumente

Im Anschluss an seine oben zitierte Rekonstruktion schreibt Keil, das Konsequenzargument beanspruche, aus allseits akzeptierten Prämissen eine Folgerung abzuleiten, die der Kompatibilist zwar ablehne, aber vernünftigerweise nicht ablehnen dürfe. „Wenn er es gleichwohl tut, scheint er eine der beiden Prämissen ablehnen zu müssen.“ (83) Diese Darstellung ist zumindest irreführend. Die von Keil diskutierte, oben zitierte Formulierung von van Inwagens Konsequenzargument ist nur dessen erste, provisorische Fassung. In seiner endgültigen Formulierung der modalen Fassung spielen zwei Schlussregeln eine zentrale Rolle, und van Inwagen betont gerade – mit den meisten zeitgenössischen Inkompatibilisten wie Kompatibilisten –, dass der schwächste Punkt des Arguments die zweite dieser beiden Regeln sei (van Inwagen 1983, 96 f.). Ich gebe im Folgenden zunächst van Inwagens Konsequenzargument mit einigen Erläuterungen wieder und stelle dann eine vereinfachte Fassung vor. Die erste der betreffenden Regeln lautet:

Regel α : Wenn p notwendig ist, dann gilt: p , und niemand hat oder hatte jemals eine Wahl in Bezug auf p .

Schreiben wir mit van Inwagen „ Np “ für „ p , und niemand hat oder hatte jemals eine Wahl in Bezug auf p “ (vgl. van Inwagen 1983, 93; ders. 2002, 156) und „ $\Box p$ “ für „ p ist logisch notwendig (wahr in jeder möglichen Welt)“, dann können wir Regel α auch wie folgt abkürzen:

$$\Box p \mid - Np$$

Diese Regel kommt in Keils Rekonstruktion nicht vor. Die zweite Regel dagegen – Regel β – kommt vor, und Keil beschreibt sie wie folgt:

„Wenn aus der Wahrheit von p notwendig die von q folgt, dann folgt aus der Notwendigkeit oder Unabänderlichkeit von p die von q . [...] Dieses Prinzip heißt in der Literatur auch ‚Regel β ‘ oder ‚Transferprinzip‘.“ (84)

Diese Darstellung von Regel β ist allerdings falsch – oder zumindest, bei wohlwollender Interpretation, sehr irreführend. Ist hier in beiden Vorkommnissen des Ausdrucks „notwendig“ beziehungsweise „Notwendigkeit“ von *logischer* Notwendigkeit die Rede? Keils Formulierung legt dies nahe. Doch in van Inwagens Regel β geht es nicht um logische Notwendigkeit, sondern um bestimmte Eigenschaften des Begriffs der mangelnden Wahlfreiheit. Und obwohl natürlich logische Notwendigkeit mangelnde Wahlfreiheit impliziert – niemand hat eine Wahl in Bezug auf logisch notwendige Sachverhalte –, sind nicht umgekehrt alle Sachverhalte, in Bezug auf die wir keine Wahlfreiheit genießen, logisch notwendig. Spricht Keil womöglich nur im Wenn-Satz von logischer, jedoch im Dann-Satz von einer anderen Art von Notwendigkeit? Auch diese Interpretation trifft auf van Inwagens Regel β nicht zu. Tatsächlich lautet diese:

Regel β : (i) Wenn p der Fall ist und niemand eine Wahl in Bezug auf p hat noch je hatte; und (ii) p q impliziert und niemand eine Wahl in Bezug auf diese Tatsache hat noch je hatte; dann gilt auch, dass niemand eine Wahl in Bezug auf q hat oder je hatte; kurz:

$Np, N(p \supset q) \vdash Nq$ (vgl. van Inwagen 1983, 93 f.; ders. 2002, 159).

Nun mag man die Eigenschaft der mangelnden Wahlfreiheit (oder der mangelnden Macht, fehlenden Kontrolle) hinsichtlich eines Sachverhalts zwar in einem sehr allgemeinen Sinne als eine Form von Notwendigkeit beschreiben. Die entscheidende Frage lautet jedoch gerade, wie sich Regel β zu der entsprechenden – in jedem so genannten „normalen“ Modalsystem gültigen – Regel verhält, in der an den Stellen, an denen in van Inwagens Regel β „N“ (für „niemand hat eine Wahl in Bezug auf“) steht, von logischer Notwendigkeit die Rede ist. Die Frage lautet, ob van Inwagens N-Operator sich wie ein „normaler“ Notwendigkeitsoperator verhält oder nicht.⁴

Ausgerüstet mit den Regeln α und β , präsentiert van Inwagen das Konsequenzargument schließlich wie folgt. P_0 stehe für eine Proposition, die einen Zustand des Universums zu einem Zeitpunkt in ferner Vergangenheit und vor der Existenz menschlicher Wesen (zum Beispiel im Präkambrium) beschreibt; L für eine Konjunktion der geltenden Naturgesetze; und P für eine beliebige wahre Proposition). Angenommen, der Determinismus sei wahr. Dann gilt zunächst:

Argument I

(1) Notwendigerweise: L und P_0 implizieren P .
 $\square (L \ \& \ P_0 \supset P)$ |Prämisse/Annahme, die aus dem Determinismus folgt.

(1) folgt aus der deterministischen These. Denn wenn der Determinismus wahr ist, dann gibt es keine mögliche Welt, in der die (in unserer Welt aktuellen) Naturgesetze herrschen sowie der Weltzustand P_0 vorlag, doch P nicht der Fall ist oder war.

(2) Niemand hat oder hatte je eine Wahl in Bezug auf den Weltzustand P_0 .
 NP_0 |Prämisse.

(3) Niemand hat oder hatte je eine Wahl in Bezug auf die Naturgesetze.
 NL |Prämisse.

(4) Notwendigerweise: Wenn P_0 , dann ist auch P der Fall, sofern L gilt.
 $\square (P_0 \supset (L \supset P))$ |1, Aussagenlogik.

(5) Niemand hat oder hatte je eine Wahl in Bezug auf die Tatsache, dass, wenn P_0 , auch P der Fall ist, sofern L gilt.
 $N(P_0 \supset (L \supset P))$ |4, Regel α .

(6) Niemand hat oder hatte je eine Wahl in Bezug auf die Tatsache, dass, wenn L gilt, P der Fall ist.
 $N(L \supset P)$ |2, 5, Regel β .

(7) Niemand hat oder hatte je eine Wahl in Bezug auf (die Wahrheit von) P .
 NP |3, 6, Regel β .

⁴ Vgl. hierzu etwa Lewis (1993), Fn. 11, und Brian Chellas' Charakterisierungen „normaler“ Modalsysteme in: Chellas (1980), Kap. 4.

Da „P“ für eine beliebige wahre Proposition steht, ergibt sich, dass niemand eine Wahl in Bezug auf irgendetwas hat.

Betrachten wir nun eine etwas leichter überschaubare Fassung des Arguments, dessen entscheidende Prämisse ebenso unproblematisch ist wie die Prämissen in van Inwagens Ur-Fassung. Man erhält jene neue, vereinfachte Fassung, wenn man (2) und (3) aus Argument I zu einer Prämisse zusammenfasst:

Argument II

- | | |
|---|--|
| (1) Notwendigerweise: Die Naturgesetze L und (der lange vergangene Weltzustand) P_0 implizieren P.
$\square(L \ \& \ P_0 \supset P)$ | Prämisse/Annahme, die aus dem Determinismus folgt. |
| (2) Niemand hat oder hatte je eine Wahl in Bezug darauf, dass L und P_0 P implizieren.
$N(L \ \& \ P_0 \supset P)$ | 1, Regel α . |
| (3) Niemand hat oder hatte je eine Wahl in Bezug auf die Wahrheit der Konjunktion L & P_0
$N(L \ \& \ P_0)$ | Prämisse. |
| (4) Niemand hat oder hatte je eine Wahl in Bezug auf (die Wahrheit von) P.
NP | 2, 3, Regel β . |

Diese vereinfachte Fassung des Arguments bietet sich an, weil – wenn man mit van Inwagen (2) und (3) aus Argument I akzeptiert – es bereits intuitiv ebenso plausibel ist anzunehmen, dass niemand eine Wahl in Bezug auf die Wahrheit einer Konjunktion hat oder je hatte, die die Naturgesetze und einen Weltzustand in ferner Vergangenheit beschreibt. Ferner lässt sich leicht zeigen, dass (3) in Argument II aus (2) und (3) in Argument I folgt, wenn Regel α und Regel β gelten.⁵ Warum van Inwagen nicht mit dieser übersichtlicheren Fassung des Konsequenzarguments arbeitet, ist mir nicht klar.

Mit Argument II haben wir ein relativ unkompliziertes Konsequenzargument in der Hand, das zeigt, dass, wenn der Determinismus wahr ist, niemand eine Wahl in Bezug auf irgendetwas hat (und also auch niemand Wahlfreiheit in Bezug auf seine eigenen Handlungen und Entscheidungen genießt). Wo, wenn überhaupt, kann dieses Argument angegriffen werden? Prämisse 1 ist wieder lediglich eine Annahme, die aus dem Determinismus folgt; (2) folgt mit Regel α – die sakrosankt ist – aus (1); (3) scheint ebenfalls nicht sinnvoll bezweifelt werden zu können, denn weder die ferne Vergangenheit noch die Naturgesetze können oder konnten je von einem menschlichen Wesen beeinflusst werden. (4) folgt klarerweise mit Regel β aus (2) und (3). Die entscheidende Frage konzentriert sich somit auf die Gültigkeit von Regel β .

Prima vista ist diese Regel enorm plausibel. Wenn niemand eine Wahl im Hinblick darauf hat oder je hatte, dass gegen Ende der Kreidezeit ein großer Meteorit auf der Erde einschlug, noch im Hinblick darauf, dass als Konsequenz dieses Ereignisses und der herrschenden Naturgesetze die letzten Dinosaurier ausstarben, dann hat und hatte offenbar auch niemand je eine Wahl hinsichtlich dessen, dass Letzteres geschah. Wenn niemand eine Wahl hinsichtlich dessen hat, dass in zehn Milliarden Jahren der Wasserstoffvorrat der Sonne aufgebraucht sein wird (wenn kein

⁵ Denn dann ist der N-Operator „agglomerativ“, das heißt, es gilt: $Np, Nq \vdash N(p \ \& \ q)$, was sich leicht wie folgt beweisen lässt: 1. Np ; 2. Nq ; 3. $\square(p \supset (q \supset (p \ \& \ q)))$ (gilt in allen modallogischen Kalkülen); 4. $N(p \supset (q \supset (p \ \& \ q)))$ (aus 3 und Regel α); 5. $N(q \supset (p \ \& \ q))$ (aus 1, 4, Regel β); 6. $N(p \ \& \ q)$ (aus 2, 5, Regel β). Vgl. hierzu McKay u. Johnson (1996).

menschliches Wesen herbeiführen kann, dass dies nicht der Fall sein wird); und wenn niemand eine Wahl in Bezug darauf hat, dass als Konsequenz dieses Ereignisses alles Leben in unserem Sonnensystem erlöschen wird, dann, so scheint es, hat auch niemand eine Wahl hinsichtlich dessen, dass dieses Ereignis stattfinden wird. Die Liste der Beispiele ließe sich beliebig verlängern. Trotz ihrer Plausibilität ist die Regel in den letzten Jahren indessen mit einigen Gegenbeispielen konfrontiert worden.⁶ Gegen diese ist ihrerseits eingewandt worden, dass sie nicht stichhaltig sind beziehungsweise dass sie keine Gefahr für leicht modifizierte β -Prinzipien darstellen, mit denen sich ebenso schlagkräftige Konsequenzargumente formulieren lassen.⁷ Im vorliegenden Rahmen kann nicht näher auf diese Diskussion eingegangen werden. Kommen wir stattdessen an dieser Stelle noch einmal auf Keils Kritik an van Inwagens Konsequenzargument zurück und werfen einen Blick auf die von ihm vorgeschlagene Alternative.

III. Ein alternatives Argument für den Inkompatibilismus?

Statt die Unabänderlichkeit der Vergangenheit zu bemühen, so moniert Keil, täte der Inkompatibilist „besser daran, klarzustellen, dass die Freiheitsgefährdung *allein* von der deterministischen Lehre ausgeht, dass niemals alternative Möglichkeiten des Weltverlaufs bestehen“ (85 f.). Wenn der Sinn des betreffenden Verbs („[einen Zustand] in der Hand haben“, „kontrollieren“, „beeinflussen“, „verändern“ oder „Anderskönnen“) im Sinne von „handelnd eine von mehreren offenen Alternativen verwirklichen“ geklärt sei, dann „wäre die Konklusion ‚Niemand hat seine Handlung in der Hand‘ nur noch eine triviale Instanz der Allaussage ‚Niemand hat je einen Weltzustand in der Hand‘“ (86). Wie wir indessen gesehen haben, ist genau das die Pointe des Konsequenzarguments: In Argument I und II steht P ja für einen Satz, der eine beliebige wahre Proposition ausdrückt, und die Überlegung führt vor, wie aus der deterministischen Lehre, dass niemals alternative Möglichkeiten des Weltverlaufs bestehen, geschlossen werden kann, dass niemand je eine Wahl in Bezug auf irgendetwas hat. Bei Licht betrachtet, so Keil weiter,

„liegt also die ganze Last des inkompatibilistischen Arguments bei der Erläuterung des zentralen Handlungsverbs. Gesucht ist ein Verb, das das Nichtbestehen alternativer Möglichkeiten der Rede über Fähigkeiten oder frei gewählte Handlungen kommensurabel macht. [...] Die metaphysische Unmöglichkeit, von der das Konsequenzargument handelt, muss der Rede von der Fähigkeit oder Unfähigkeit, frei seine Handlungen auszuführen, kommensurabel gemacht werden, um den Kompatibilisten zu überzeugen. Als Mittelbegriffe bieten sich die Verben ‚wählen‘ oder ‚entscheiden‘ an.“ (86)

Hier beschreibt Keil jedoch ebenfalls genau das, was van Inwagens Konsequenzargument tut. Er präsentiert es als eine Überlegung, die, wie wir gesehen haben, zeigt, wie aus der aus dem Determinismus folgenden kausal-naturgesetzlichen Unmöglichkeit, dass jemals etwas anders ist, als es ist, folgt, dass niemand sich je anders entscheiden kann, als er/sie es tatsächlich tut.

Abschließend schlägt Keil eine alternative inkompatibilistische Überlegung im Geiste des Konsequenzarguments vor. Er fährt fort:

⁶ Die meistdiskutierten Kritiken stammen von: Slotte (1982), Widerker (1987), McKay u. Johnson (1996). Slotte bezieht sich auf van Inwagen (1975).

⁷ Vgl. hierzu insbesondere O'Connor (1993), Crisp u. Warfield (2000), van Inwagen (2002).

„Der Inkompatibilist könnte dann argumentieren: (1) Wenn der Determinismus wahr ist, bestehen niemals alternative Möglichkeiten des Weiterverlaufs. (2) Wenn niemals alternative Möglichkeiten des Weiterverlaufs bestehen, gibt es auch niemals etwas zu wählen oder zu entscheiden. (3) Also entscheiden wir auch nicht über die Ausführung unserer Handlungen. Im Unterschied zu van Inwagens Argument wird hier nicht auf die Vergangenheit Bezug genommen.“ (86)

Was die Konklusion (3) angeht, so mag man anmerken, dass der Inkompatibilist typischerweise bestreitet, dass ein vollständig determiniertes Verhalten in einem hinreichend robusten Sinne als „Handlung“ bezeichnet werden kann. Keil selbst betont dies an verschiedenen Stellen seines Buches: „Handelnkönnen“, so schreibt er, „*heißt* schon So-oder-anders-Können. Das Anderskönnen ist, technisch ausgedrückt, eine *analytische Komponente* des Handelnkönnens.“ (89, vgl. auch 118) Nimmt man dies ernst, so sollte man in (3) statt von „Handlungen“ besser von „Verhalten“ sprechen oder von Geschehnissen, die unter nicht-deterministischen Bedingungen als „Handlungen“ bezeichnet werden könnten. Wichtiger ist das folgende Problem.

Keil beruft sich in seinem obigen Argument zwar in der Tat nicht *explizit* auf die Vergangenheit. Doch bezieht er sich (mehr oder weniger implizit) auf die Zukunft. Der Vorteil dieses Manövers bleibt, auch im Lichte von Keils Kritik an van Inwagens Bemühung der Vergangenheit, unklar. Denn dass es, wenn der Determinismus wahr ist, „niemals alternative Möglichkeiten des Weiterlaufs“ gibt, bedeutet, dass unter Annahme des Determinismus für beliebige Zeitpunkte t feststeht, wie die Welt zu einem beliebigen von t aus gesehen späteren Zeitpunkt aussehen wird. Da t aber zu jedem von t aus gesehen späteren Zeitpunkt vergangen ist, läuft Keils im Rahmen seiner Rede von „Weiterverlauf“ auf jeweils zukünftige Zustände bezogene Aussage inhaltlich auf dasselbe hinaus wie die von ihm kritisierte These, dass, wenn der Determinismus wahr ist, zu jedem Zeitpunkt des Weltverlaufs der dann gegenwärtige Gesamtzustand der Welt durch die Naturgesetze und einen beliebigen vergangenen Gesamtzustand der Welt festgelegt ist. Keil treibt den vermeintlichen Teufel mit dem Beelzebub aus.

Drittens wirft Keil jedoch mit seinem Alternativvorschlag eine Frage auf, der sich der Konsequenztheoretiker in der Tat stellen sollte: Wozu eigentlich bedarf der Inkompatibilist des (einigermaßen komplexen) Konsequenzarguments im Stile von van Inwagens und seinen Mitstreitern? Bringt nicht eine schlichtere Überlegung, wie Keil sie präsentiert, den Kern der inkompatibilistischen Idee ebenso gut auf den Punkt? Wenn der universale kausale Determinismus wahr ist, dann gibt es niemals Alternativen zum tatsächlichen Weltverlauf. Und wenn das der Fall ist, dann hat trivialerweise auch niemand jemals eine Wahl zwischen verschiedenen Alternativen. Daher genießt auch niemand Wahl- oder Entscheidungsfreiheit hinsichtlich der Entscheidungen, die er/sie trifft, und des mit diesen Entscheidungen verknüpften Verhaltens.

Die Antwort lautet, dass van Inwagens Konsequenzargument genau diese Überlegung genauer analysiert, ihre Prämissen explizit macht und ihre, wie man sagen könnte, Tiefengrammatik an den Tag bringt. Was genau bedeutet kausal-naturgesetzlicher Determinismus und *warum* ergibt sich aus ihm, dass es zu keinem Zeitpunkt alternative Möglichkeiten des Weiterverlaufs gibt? Weil dann notwendigerweise gilt, dass eine Beschreibung der Naturgesetze und eines beliebigen vergangenen Gesamtzustands der Welt jede Beschreibung eines beliebigen späteren Weltzustands implizieren. Warum folgt hieraus, dass es niemals etwas zu entscheiden gibt? Weil es niemals etwas hinsichtlich der Naturgesetze und vergangener Weltzustände zu entscheiden gibt. Die obigen Argumente I und II machen dies deutlich.

IV. Wahlfreiheit und Verantwortlichkeit

Keil setzt sich im dritten Kapitel seines Buches auch mit dem so genannten Prinzip der alternativen Möglichkeiten auseinander (Abschnitt 3.4). Er diskutiert dort Harry Frankfurts berühmte Kritik an diesem Prinzip (Frankfurt 1969) sowie die wichtigsten Einwände gegen Frankfurts Angriff. Auf diese Diskussion kann ich hier nicht mehr eingehen; doch sei abschließend skizziert, welche Rolle das Prinzip der alternativen Möglichkeiten meines Erachtens für das Konsequenzargument spielt.

Die bisher betrachteten inkompatibilistischen Argumente arbeiten mit Begriffen wie Wahl- und Entscheidungsfreiheit. Ihre Konklusion behauptet, was man als *Wahlfreiheits-Inkompatibilismus* bezeichnen kann: Der Determinismus, so besagt sie, ist unvereinbar mit der Annahme, dass man je Alternativen zu dem hat, was man faktisch tut. Die entscheidende Frage jedoch, um die es dem typischen Willensfreiheits- und Wahlfreiheitstheoretiker am Ende geht, lautet, ob und, wenn ja, unter welchen Bedingungen Personen für ihre Handlungen und Entscheidungen *moralisch verantwortlich* sind. Der klassische Inkompatibilist geht in diesem Zusammenhang üblicherweise ohne viel Aufhebens von einem Wahlfreiheits-Inkompatibilismus zu einem *Verantwortlichkeits-Inkompatibilismus* über: zu der These, dass Determinismus und moralische Verantwortlichkeit nicht miteinander vereinbar sind. Wodurch ist dieser Schritt gerechtfertigt? Hier kommt das Prinzip der alternativen Möglichkeiten ins Spiel. In einer typischen Formulierung besagt es, dass eine Person nur dann für eine von ihr ausgeübte Handlung moralisch verantwortlich ist, wenn sie auch anders hätte handeln können.

Diese Behauptung verbindet den Wahlfreiheits-Inkompatibilismus mit dem Verantwortlichkeits-Inkompatibilismus, denn der erstgenannte behauptet ja, dass wir im Falle des Determinismus niemals anders hätten handeln und entscheiden können, als es tatsächlich geschieht. Und wenn das so ist, ist die Bedingung, die das Prinzip der alternativen Möglichkeiten für moralische Verantwortlichkeit aufstellt, nicht erfüllt. Keil verteidigt das Prinzip gegen die Einwände Frankfurts, bringt es jedoch nicht unmittelbar mit dem Konsequenzargument in Verbindung. Gleichwohl gilt: Wenn eine Verteidigung des Prinzips der alternativen Möglichkeiten gelingt, aber nur dann, kann das Wahlfreiheits-Konsequenzargument im Nachgang zu seiner Konklusion auch eine entsprechende Schlussfolgerung in Bezug auf Verantwortlichkeit ziehen und folgern: Der Determinismus ist mit der Existenz moralischer Verantwortlichkeit nicht vereinbar. Steht und fällt die entscheidende inkompatibilistische These – der Verantwortlichkeits-Inkompatibilismus – somit mit der Gültigkeit des Prinzips der alternativen Möglichkeiten?

Keineswegs. Denn es gibt eine Form des Konsequenzarguments, die *direkt*, das heißt ohne einen Umweg über Begriffe wie Wahl- und Entscheidungsfreiheit, Kontrolle usw. und somit ohne den Rückgriff auf das Prinzip der alternativen Möglichkeiten zur These des Verantwortlichkeits-Inkompatibilismus führt (vgl. van Inwagen 1983, 183–189). In Argument II ersetze man hierzu die Rede von „niemand hat oder hatte je eine Wahl in Bezug auf/darauf, dass“ in (2), (3) und der Konklusion (4) durch „niemand ist oder war jemals moralisch verantwortlich für/dafür, dass“ beziehungsweise interpretiere den N-Operator in der formalen Fassung des Arguments in diesem letztgenannten Sinne. Die Prämissen des Arguments bleiben akzeptabel und die Schlüsse gültig. Wieder konzentriert sich die entscheidende Frage daher auf das Regel β entsprechende Transfer- oder Geschlossenheitsprinzip:

Regel β_v : (i) Wenn p der Fall ist und niemand *moralisch verantwortlich* für p ist oder je war, und (ii) p q impliziert und niemand für diesen Sachverhalt moralisch verantwortlich ist oder je war, dann gilt (iii), dass auch niemand je moralisch verantwortlich für q ist oder war.

Auch dieses Prinzip ist *prima vista* plausibel; doch auch dieses Prinzip ist in den letzten Jahren kontrovers diskutiert worden.⁸ Insbesondere Überdeterminierungseinwände haben ihm zugesetzt: Angenommen, ein gegebener Sachverhalt Q folgt nicht nur aus dem Sachverhalt P , für den niemand verantwortlich sei, sondern zusätzlich aus einem anderen Sachverhalt, für den sehr wohl jemand verantwortlich sei. Dann, so hat man argumentiert, ist sehr wohl jemand für Q verantwortlich, auch wenn weder jemand für P noch für $P \supset Q$ verantwortlich ist. Wenn diese Überlegung korrekt ist, liegt ein Gegenbeispiel gegen Regel β_v vor. Die Aussichten von Konsequenzargumenten für einen *moderaten Verantwortlichkeits-Inkompatibilismus* – einen Inkompatibilismus, der behauptet, dass, wenn der Determinismus wahr ist, kein Akteur in der aktuellen Welt moralisch für etwas verantwortlich ist – hängen nicht davon ab, ob die Prämissen dieser Argumente sich auf die ferne Vergangenheit berufen, sondern davon, ob sie sich zum Beispiel gegen Überdeterminierungseinwände der gerade skizzierten Art verteidigen lassen.⁹

Prof. Dr. Christoph Jäger, University of Aberdeen, Department of Philosophy; Universität Innsbruck, Katholisch-Theologische Fakultät, Institut für Christliche Philosophie, Karl-Rahner-Platz 1, 6020 Innsbruck, Österreich

Literatur

- Campbell, Joseph Keim (2007), Free Will and the Necessity of the Past, in: *Analysis*, 67, 105–111.
- Chellas, Brian (1980), *Modal Logic – An Introduction*, London.
- Crisp, Thomas u. Warfield, Ted (2000), The Irrelevance of Indeterministic Counterexamples to Principle Beta, in: *Philosophy and Phenomenological Research*, 61, 173–184.
- Fischer, John Martin u. Ravizza, Mark, S. J. (1998), *Responsibility and Control – A Theory of Moral Responsibility*, Cambridge.
- Frankfurt, Harry G. (1969), Alternate Possibilities and Moral Responsibility, in: *The Journal of Philosophy*, 66, wieder abgedruckt in: ders., *The Importance of What We Care About*, Cambridge 1994, 1–10.
- Ginet, Carl (1966), Might We Have No Choice?, in: *Freedom and Determinism*, hg. v. K. Lehrer, New York, 87–104.
- Ginet, Carl (1990), *On Action*, Cambridge.
- Jäger, Christoph (2006), Drei Konsequenzargumente für eine inkompatibilistische Theorie moralischer Verantwortung, in: *Zeitschrift für philosophische Forschung*, 60, 504–527.
- Kane, Robert (1996), *The Significance of Free Will*, Oxford.
- Kapitan, Tomis (2002), A Master Argument for Incompatibilism?, in: *The Oxford Handbook of Free Will*, hg. v. R. Kane, Oxford, 127–157.
- Lewis, David (1993), Evil for Freedom's Sake?, in: *Philosophical Papers*, 22, deutsch in: *Analytische Religionsphilosophie*, hg. v. Ch. Jäger, Paderborn 1998, 273–301.
- McKay, Thomas u. Johnson, David (1996), A Reconsideration of an Argument against Compatibilism, in: *Philosophical Topics*, 24, 113–122.

⁸ Vgl. hierzu etwa Ravizza (1994), Warfield (1996), Stump u. Fischer (2000).

⁹ Ich habe dies in Jäger (2006) getan, indem ich eine leicht modifizierte Version von Regel β_v vorschlage.

- O'Connor, Timothy (1993), On the Transfer of Necessity, in: *Nous*, 27, 204–218.
- Ravizza, Mark (1994), Semi-Compatibilism and the Transfer of Nonresponsibility, in: *Philosophical Studies*, 75, 61–93.
- Slote, Michael (1982), Selective Necessity and the Free-will Problem, in: *The Journal of Philosophy*, 79, 5–24.
- Stump, Eleonore u. Fischer, John Martin (2000), Transfer Principles and Moral Responsibility, in: *Philosophical Perspectives*, 14, Oxford, 47–55.
- Van Inwagen, Peter (1975), The Incompatibility of Free Will and Determinism, in: *Philosophical Studies*, 27, 185–199.
- Ders. (1983), *An Essay on Free Will*, Oxford.
- Ders. (1989), When is the Will Free?, in: *Philosophical Perspectives*, 3, 1989, zitiert nach dem Wiederabdruck in: *Agent, Causes, and Events*, hg. v. T. O'Connor, Oxford, 219–269.
- Ders. (2002), Free Will Remains a Mystery, in: *The Oxford Handbook of Free Will*, hg. v. R. Kane, Oxford, 158–177.
- Warfield, Ted (1996), Determinism and Moral Responsibility Are Incompatible, in: *Philosophical Topics*, 24, 215–226.
- Warfield, Ted (2000), Causal Determinism and Human Freedom are Incompatible: A New Argument for Incompatibilism, in: *Philosophical Perspectives*, 14, 167–180.
- Widerker, David (1987), On an Argument for Incompatibilism, in: *Analysis*, 47, 37–41.
- Wiggins, David (1973), Towards a Reasonable Libertarianism, in: *Essays on Freedom of Action*, London, 31–61.
- Willaschek, Marcus (2008), Freiheit und Vernunft – Weshalb der Inkompatibilismus auf einer absolutistischen Konzeption von Vernunft beruht, in: *Philosophisches Jahrbuch*, 115, 2. Halbband, 397–417.

Abstract

In his recent book *Willensfreiheit* (2007) Geert Keil defends a version of libertarianism. Yet he criticizes a flagship argument for incompatibilism. Van Inwagen's consequence argument, Keil thinks, relies on an irrelevant premise when it claims that agents have no choice about the remote past. I argue that Keil's charge rests on a misunderstanding. I then sketch why discussions of the consequence argument should focus on the question whether or not a certain version of rule Beta is valid.